SECRETOR STATUS, *FUCOSYLTRANSFERASE* 2 *(FUT2)* GENE POLYMORPHISMS AND SUSCEPTIBILITY TO HIV INFECTIONS AMONG FEMALE SEX WORKERS IN NAIROBI, KENYA

Nadia Musimbi Chanzu

W80/83581/2012

A Thesis Submitted in Fulfillment for the Degree of Doctor of Philosophy (Ph.D) in the University of Nairobi Institute of Tropical and Infectious Diseases

University of Nairobi

2014

DECLARATION

This Thesis is my original work and has not been presented for a degree in any other University.

Nadia Musimbi Chanzu

Signed: ……

This Thesis has been presented with our approval as University Supervisors:

Prof. Walter Mwanda, MBChB., MD. Signed: ……

Date: ……

Prof. Omu Anzala, MBChB, PhD. Signed: ……

Date: ……

Dr. Julius Oyugi, MSc. Ph.D Signed: ……

Date: ……

DEDICATION

This Thesis is dedicated to my parents, Yusuf Chanzu and Jamila Chanzu, my siblings, Abdalla, Issa, and Zahra. Thank you for the confidence you have always had in me. Thank you for your tremendous encouragement, support and prayers.

I also dedicate this book to my late grandfather Abdalla Baraka Chanzu and my late great grandfather Baraka Lwoya Chanzu for their foresightedness which encouraged our family acquire high level quality education.

ACKNOWLEDGEMENTS

This work has had the support and goodwill of so many people that it is difficult to acknowledge each one of them individually,

I would like to thank my entire family for their continued support. A special mention to my mother Jamila Chanzu, for always encouraging me to pursue my dreams, and teaching me that education was truly the key to success, her patience and kind heart. My father, Yusuf Chanzu for his continued support, mentorship and for teaching me that there is nothing impossible. He taught and led by example, and supported me through every success and every failure from the time I was a child. Above all, he taught me the importance of being persistent and consistent.

I would also like to thank my supervisors, Professor Walter Mwanda, for encouraging me right from inception, generating it into a concept proposal and eventually ensuring the research work was completed. Prof. Mwanda was always firm, but patient, understanding, and encouraged me all through. I would also like to thank Prof. Mwanda for being my mentor throughout my scholarship on the CIHR, International Infectious Minds and Global Health Training Program (CIHR, IID & GHTP).

My thanks also go to Professor Omu Anzala for supporting me towards securing the grant funding to finance this entire project and for supporting my visits to conferences both regionally and internationally, and through all seminars, where I was required to disseminate my results. Above all I thank Prof. Anzala for his patience, and

encouragement and for being my mentor as I served in the capacity of a Project Leader managing the grant funding from the International Development Research Centre (IDRC).

I also thank Dr. Julius Oyugi for having confidence and supporting me right from the beginning, providing me an opportunity to grow as a leader in coordinating a number of activities as I carried out my research project. I would particularly like to thank him for his encouragement as I coordinated Journal Club sessions for two years. This widened my scope, allowed me the opportunity to learn across diverse fields and grow as a person. I would especially like to thank Dr. Oyugi for granting me an opportunity to be part of the International Infectious Diseases and Global Health Training Program (IID & GHTP), Kenyan trainees' team. This was an excellent scholarship, and learning opportunity.

Further I thank, Professor Keith Fowke, for awarding me an international scholarship on the prestigious CIHR, International Infectious Diseases and Global Health Training Program (IID & GHTP). This program provided me with the real global experience in health research. The program supported all my international learning visits, and allowed me the opportunity to gain a deep appreciation of science especially infectious disease based research all around the world. I would particularly like to thank Prof. Fowke for always encouraging me during all my presentations, and pushing me to aim higher.

I would like to thank Dr. Joshua Kimani for always having an open door, a listening ear and the endless doses of encouragement. He supported me through my initial project

presentations, ethical preparations and through my entire project.

I would like to thank the entire Kenya AIDS Control Project team. This study was coordinated from the SWOP clinics to the UNITID laboratories and to KAVI-ICR laboratories. A special mention to Wendy, Marion, Anne of KACP and Jackton and IreneAnn of KAVI-ICR. I would like to thank all SWOP Clinic staff, who made sure this project was a real success from inception. A special mention to Njoki, Martin and Dr. Makobu from the three clinics.

I would like to thank all KAVI-ICR staff, who supported my project from the pilot stage to completion. A special mention to Bashir Farah the laboratory manager at KAVI-CR, for ensuring all my laboratory reagents were procured, all protocols were optimized and successfully implemented and always pushing me to move forward.

I would also like to thank Nancy Thairu, Charles Osemba both of KAVI-ICR and Dinah Amwayi of the Kenya AIDS Control Project for ensuring all the financial aspects of the project were met, and supporting me towards grant funds management.

I am indebted to Moses Muriuki - KAVI-ICR, for my database design. A special thank you to Mr. Wycliffe Ayieko of the University of Nairobi for the comprehensive data management and to Dr. George Michuki of International Livestock Research Institute (ILRI) for the sequencing data management

I also wish to appreciate, Professor Christine Kigondu. She has offered me motherly advice, been my friend, my mentor and my greatest strength. Prof. Kigondu was my supervisor during my undergraduate project, where I discovered my passion for research and attained a straight A under her mentorship. During my masters she encouraged me despite the distance, and yet again under her mentorship attained an outstanding scholarship to pursue my project. On returning back home, Prof. Kigondu encouraged me all through to ensure I would one day graduate with a doctorate degree. Thank you for holding my hand all through.

Last but not least, I am grateful to the Global Health Research Initiative/International Development Research Centre for awarding me grant funding towards this research project. I would also like to appreciate the Canadian Institutes of Health Research, International Infectious Diseases and Global Health Training Program for awarding me an international scholarship as I pursued my studies.

Above all, this work is a blessing from the Almighty God, truly an answered prayer from above.

TABLE OF CONTENTS

TABLE OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ABO – A system of classifying blood groups based on the A or B antigen expression

AIDS – Acquired Immunodeficiency Syndrome

ART – Antiretroviral Therapy

ATP – Adenosine Tri-Phosphate

BV – *Bacterial vaginosis*

CA - Capsid

CBD – Central Business District

CC – Cysteine-Cysteine

CCDS – Charged Couple Devices

CCL3 – CC Chemokine Ligand 3

CCR5 – CC Chemokine Receptor 5

CD – Cluster of Differentiation

cDNA – Complementary DNA

CDS – Protein Coding Sequences

CI – Confidence Internal

CIHR, IID & GHTP – Canadian Institutes of Health Research, International Infectious Diseases and Global Health Training Program

CRFs – Circulating Recombinant Forms

CXC – Cysteine-amino acid-Cysteine

DARC – Duffy Antigen Receptor for Chemokines

DC – Dendritic Cell

DNA – Deoxyribonucleic Acid

dsDNA – Double stranded Deoxyribonucleic Acid

EC – Elite Controllers

ECD – Extracellular Domain

EDTA – Ethylenediaminetetraacetic acid

FSW – Female Sex Worker

FUT1 – Gene encoding for Enzyme Fucosyltransferase 1

FUT2 – Gene encoding for Enzyme Fucosyltransferase 2

FY – Gene encoding for Duffy blood group antigen

GalNAc – N-Acetylgalactosamine

GALT – Gut-Associated Lymphoid Tissue

GC – Gonorrhea Culture

GCLP – Good Clinical Laboratory Practice

Gp - Glycoprotein

GPI - Glycosylphosphatidylinositol

GSLs – Glycosphingolipids

HESN – Highly Exposed Sero-Negatives

HIV – Human Immunodeficiency Virus

HLA – Human Leukocyte Antigen

HTLV-1 – Human T-Lymphotropic Virus - 1

IFN – Interferon

Ig – Immunoglobulin

IL-8 – Interleukin 8

ISBT – International Society of Blood Transfusion

KAIS – Kenya AIDS Indicator Survey

KAVI-ICR – KAVI Institute of Clinical Research

KDHS – Kenya Demographic and Health Survey

LC – Langerhans Cells

LTNPs – Long Term Non Progressors

MA - Matrix

MARPs- Most At Risk Populations

MEGA – Molecular Evolutionary Genetic Analysis

MIP – Macrophage Inflammatory Protein

MoT – Mode of Transmission

mRNA – Messenger Ribonucleic Acid

MSM – Men who have Sex with Men

NACC- National AIDS Control Council

NASCOP – Kenya National AIDS & STI Control Programme

NBTS – Kenya National Blood Transfusion Services

NEC – Necreotising Enterocolitis

OR – Odds Ratio

PBMCs – Peripheral Blood Mononuclear Cells

PBS – Phosphate Buffered Saline

PCR – Polymerase Chain Reaction

PIC –Pre-Integration Complex

PLHA – People Living with HIV/AIDS

PPi – Pyrophosphate

PUDs – People who Use Drugs

RANTES – Regulated upon Activation, Normal T-cell Expressed, and Secreted

RBC – Red Blood Cell

Rh – Rhesus

RNA – Ribonucleic Acid

RT – Reverse Transcriptase Enzyme

SD – Standard Deviation

SDF-1 – Stromal Cell-Derived Factor 1

SIV – Simian Immunodeficiency Virus

SNPs - Single Nucleotide Polymorphisms

SQA – Sequencing Analysis

SSA- Sub Saharan Africa

STI – Sexually Transmitted Infection

SWOP – Sex Worker Outreach Program

TM – Transmembrane

TTI – Transfusion Transmissible Infections

TV – *Trichomonas vaginalis*

UNGASS – United Nations General Assembly Special Session on HIV and AIDS

UNITID – University of Nairobi Institute of Tropical and Infectious Diseases

VBD – Voluntary Blood Donors

VLPs – Virus-Like Particles

ABSTRACT

Background:

Blood group antigens are expressed on red blood cells however; these antigens can also be expressed on some other cells particularly on the surface of epithelial cells and may be found in mucosal secretions. The gene known to determine the secretion of these blood group antigens is the *Secretor Fucosyltransferase 2* (*FUT2*) gene. In many human populations 80% secrete ABO antigens (termed secretors) while 20% do not (termed non-secretors). Furthermore, there are disease conditions that are associated with secretor status. It is against this background that this study was proposed.

Hypothesis:

There are correlations between *Secretor FUT2* gene polymorphisms and blood group antigen secretor status; non-secretors are less susceptible to HIV infections.

Broad Objective:

To investigate associations between mucosal blood group antigen expression profiles (secretor status), *Secretor FUT2* gene polymorphisms and susceptibility to HIV infection among female sex workers in Nairobi, Kenya.

Study Design:

This was a cross-sectional study.

Study Area:

This study was conducted at the Nairobi Regional Blood Transfusion Centre, and at sex worker outreach program clinics, from the Pumwani Majengo female sex worker cohort, Nairobi, Kenya.

Study Population:

This study enlisted 142 adults of both gender (male and female, aged 18-65 years) from the regional blood transfusion centre in Nairobi, Kenya. In addition, this study recruited 280 female sex workers from the well-established Pumwani Majengo cohort aged 18 to 65 years of age (n=422).

Materials and Methods:

Blood, saliva and female genital tract (vaginal and cervical) specimens were collected from each study participant once informed written consent was obtained. The laboratory analyses were carried out at the KAVI Institute of Clinical Research (KAVI-ICR) and the University of Nairobi Institute of Tropical and Infectious Diseases (UNITID) laboratories, Nairobi. Blood typing was determined using standard serological techniques using monoclonal antibodies to the ABH, Rhesus (D) and Duffy (Fya, Fyb) blood group antigens. Secretor phenotyping was determined using lectins specific to blood group H antigen in both salivary and female genital tract samples. This was correlated to the HIV sero-status. The correlation of secretor phenotypes to CD4+ T cell counts, was based on retrospective data analyses, following immunophenotyping using a CD3/CD4/CD45 panel on an BD LSR II flow cytometer. For secretor genotyping, DNA was extracted

from frozen whole blood samples. Quantitative real-time amplification was performed based on the polymerase chain reaction (PCR) on a Rotor-gene Q (Qiagen) and PCR products sequenced on the next-generation pyrosequencer, Pyromark Q24 platform (Qiagen).

Results:

Objective 1: ABO blood group phenotype frequency distribution was O>A>B>AB comprising 199 (47.2%), 120 (28.4%), 84 (19.9%) and 19 (4.5%) respectively with 408 (96.7%) Rhesus (D) positive cases. Duffy positive phenotypes were reported in 2 (0.47%) of the study participants.

Objective 2: Saliva testing showed that among the blood donors, 121 (85%) were secretors and 21 (15%) were non-secretors, while among the female sex workers, 212 (76%) were secretors and 68 (24%) non-secretors

Objective 3: Based on HIV screening, 92 (32.9%) of the female sex workers were HIV-1 infected and 188 (67.1%) HIV-1 uninfected. There was a correlation between HIV infections and secretor phenotypes. The proportion of secretors was significantly higher among women with HIV infection (77/92 = 83.7%) in comparison to HIV un-infected women (135/188 = 71.8%) ($p=0.029$). Furthermore, the incidence of HIV infection was significantly higher among blood group A secretors ($p=0.028$) in comparison to O secretors, but not B and AB.

Objective 4: The correlation to CD4+ T cell counts demonstrated, although the secretors were more susceptible to HIV, following infection, the ABH secretors maintained elevated levels of CD4+ T cell counts in comparison to non-secretors and this difference

was statistically significant (p=0.005).

Objective 5: *Secretor FUT2* genotyping demonstrated the secretor status of the population was not dependent on the allele at position 428 as has been previously reported in African populations, suggesting the secretor and non-secretor phenotype profiles may be due to novel polymorphisms in the Kenyan population.

Conclusion:

The prevalence of ABO, Rhesus and secretor/non-secretor phenotypes are similar to the global profiles as previously reported across African populations. However, the underlying genotypic variation resulting in the non-secretor phenotype in a subset of the Kenyan population as screened among our study cases does not fall under a the 'se^{428} umbrella' of a restricted geographical distribution as has been previously reported. These findings further suggest the non-secretor phenotype may confer a certain degree of protection against HIV infection, as there were higher HIV infection incidence rates among ABH secretors; but following infection, the same population of secretors maintained elevated CD4+ T cell counts.

Recommendations:

1. It is important to know the frequencies of various blood group antigen phenotypes in a population. This information is important to confirm the prevalence of both major and minor blood group phenotypes in a population. There are a number of blood group phenotypes, which have been associated to varying disease/infection susceptibilities as seen in this study, blood group A individuals were at an increased

risk for HIV infections.

2. Secretor and non-secretor phenotype profiles in populations should be determined. This data serves as a baseline for investigative studies on the correlation between secretor and non-secretor phenotypes and associations to disease conditions. In the Kenyan population, secretors were found to have a significantly higher incidence of HIV infection rates.
3. HIV awareness programmes, should include information on secretor status testing as individuals who are ABH secretors are at an increased risk of HIV infections.
4. HIV treatment and care packages, should include secretor testing as non-secretor individuals have been shown to progress faster to infection, based on a rapid decline in CD4+ T cell counts in comparison to the secretor counterparts.
5. Secretor testing should be comprehensive, based on both phenotyping and genotyping. There may be silent and/or novel mutations specific to a population, which confer varying phenotypic traits, hence dual screening will ensure an individual is ascertained of their secretor status.

1 INTRODUCTION

Thirty years since the discovery of the Human Immunodeficiency Virus (HIV), the etiologic agent of Acquired Immunodeficiency Syndrome (AIDS), the overall growth of this global epidemic appears to have stabilized (UNAIDS, 2013). Moreover, the past decade has witnessed a significant decline in AIDS-related deaths, owing to the scale up in Antiretroviral Therapy (ART). Despite these advances, the pandemic remains a global emergency, as new HIV infections continue to accumulate annually. UNAIDS (2013) reports nearly 35 million people are living with HIV/AIDS, with 2.3 million new infections every year (this translates to 6,000 new infections per day). Sub-Saharan Africa is the hardest hit, bearing an estimated two-thirds of the burden, with 1.6 million Kenyans living with HIV (KAIS, 2012).

Kenya faces a mixed HIV/AIDS epidemic that continues to exert a devastating toll on all sectors of society. However, as it is globally, the overall prevalence of HIV appears to have stabilized at 5.6% (KAIS, 2012). Furthermore, transmission of HIV in Kenya is primarily through heterosexual contact, which accounts for 93.7% of new infections, and the epidemic disproportionately affects young women. It is estimated that of the 101,560 new infections witnessed in 2013, 50% (50,530) were new infections in women, compared to the 38,000 new infections in men. In addition, the most at risk populations (MARPs) include the Female Sex Workers (FSWs) and their clients, who account for 14.1% of new infections (MoT, 2008). This study therefore investigated HIV transmission among a cohort of female sex workers. The concentrated epidemics among this high-risk population are known to have a significant impact on the HIV prevalence dynamics.

Susceptibility to HIV infection is variable within a population, and is governed by numerous factors including host genetics and mucosal immunology. In line with this, there are currently a number of efforts all geared towards curbing the HIV/AIDS situation globally. Most recently the focus has now moved to prevention measures such as vaccines and microbicide candidates that target the earliest stage of infection; preventing the establishment and expansion of small founder virion populations (Hladik *et al.*, 2007). This is based on the premise that the probable most effective prevention strategies should be mucosal, averting systemic infection and/or ensuing pathologies. Based on this, this research work was designed to focus on the female genital tract mucosa. The main aim was to profile the variable expression of ABO blood group antigens and their effect on the HIV infectious process, which has not been explored before in Kenya.

There is now extensive knowledge of the carbohydrate chemistry, enzymology, molecular genetics, and structural and evolutionary biology of the ABO blood group system (Giri *et al.*, 2011; Storry and Olsson, 2009). To date, this major blood group system remains the most important clinically in transfusion and transplantation medicine and, despite its relative simplicity; the ABO blood group system has important biological implications, dividing the world's population including both patients and donors into four major groups, A, B, AB and O. This study began with a baseline profile of these phenotypes in the Kenyan population.

The precursor substrate for both A and B antigens is the H antigen. Based on the expression profile, the H antigen is synthesized by one of two fucosyltransferases encoded by two closely related genes, *fucosyltransferase 1* (*FUT1*) and *secretor fucosyltransferase* (*FUT2*). *FUT1* encodes for a fucosyltransferase mainly responsible for

the synthesis of the H antigen on carbohydrate precursors found on red blood cells (Koda *et al.*, 2001; Ferrer-Admetlla *et al.*, 2009). While *FUT2,* the *secretor fucosyltransferase 2* gene, encodes for the enzyme alpha 1,2 fucosyltransferase and accounts for the expression of the H antigen and thereby A and B antigens, expressed in epithelial cells, body tissues and in mucosal secretions. Based on the variable expression of the *FUT2* gene, approximately 80% of the population are 'secretors (*Se*)' and 20% are 'non-secretors (*se*)' (Ferrer-Admetlla *et al.*, 2009). In secretors, the *FUT2* gene is expressed in mucin-secreting cells of various mucous membranes and glands, resulting in the secretion of the corresponding blood group antigens in body fluids and secretions. While, in non-secretors, blood group antigen expression is restricted to the erythrocyte membrane. This study therefore profiled the secretor and non-secretor phenotypes in the Kenyan population, which has not been done before.

Mutations in the second exon of the *FUT2* gene often result in a non-secretor phenotype. However, among the secretors, there are both homozygous secretors (*SeSe*) and heterozygous secretors (*Sese*). The latter, are thought to secrete less amounts of the protein (enzyme: alpha 1,2 fucosyltransferase) in comparison to the counterpart homozygous secretors. To date, over 20 single nucleotide polymorphisms have been described across the *FUT2* gene, resulting in various allelic variants of the secretor and non-secretor phenotypes (Koda *et al.*, 2001; Birney *et al.*, 2007; Ferrer-Admetlla *et al.*, 2009). A large number of these resultant polymorphisms of the gene are population specific and the most common SNP associated with the non-secretor phenotype in African populations is the non-functional *FUT2 se428,* which results in a stop codon and thereby production of a non-functional form of the enzyme alpha (1,2) fucosyltransferase

(Soejima *et al.*, 2007). However, this has not been confirmed in the Kenyan population. This study screened for the non-functional *FUT2 se428* null allele in the Kenyan population.

The *Secretor FUT2* gene and the resultant phenotypes are of particular interest, owing to the fact that there are now well-established correlations between secretor status, *FUT2* gene polymorphisms, and susceptibility to various bacterial and viral infections (Thorven *et al*., 2005; Linden *et al*., 2008; Higgins *et al*., 2009). However, few large studies have determined how secretor status and *FUT2* genetics impact HIV susceptibility. In one study of Senegalese commercial sex workers, non-secretors were found to have a reduced HIV-1 seroprevalence (Ali *et al.*, 2000). In a second study, Kindberg *et al.*, (2006) reported a slower HIV-1 disease progression in non-secretors. This study was therefore designed to investigate correlations between blood group antigen expression profiles, and their impact on HIV infections, based on the hypothesis that there are distinct correlations between secretor status and susceptibility to HIV infections in the Kenyan population; non-secretors are less susceptible to HIV infections.

In addition to the above, the CD4+ T cells are known as the hallmark of disease progression markers. This study therefore sought to investigate the role of secretor phenotypes in the context of HIV infections, based on CD4+ T cell counts as a surrogate marker of disease progression. This multi-faceted approach, allowed the comprehension of the impact of both red blood cells and T cells, on HIV infection rates.

Kenya is an ideal setting to conduct HIV research relevant to sub-Saharan Africa, where the challenges of the HIV pandemic are more acute than in other parts of the world. Sex workers have multiple sex partners and co-factors for HIV acquisition, placing them at

increased risk of HIV infections. This was therefore a very relevant study population to address these research questions. If any protective associations are observed with blood group antigen expression profiles, understanding of the mucosal mechanism of protection may provide insight into the development of new HIV preventive technologies.

1.1 Rationale and Justification

This study was designed to address a gap in HIV research, which has not been explored before in Kenya; integration of a multifaceted approach to contribute towards the elucidation of the establishment of HIV infections. The study evaluated the impact of both red cells and a specific T cell subset on HIV infections based on the hypothesis, non-secretors are less susceptible to HIV infections. The study began with a baseline profile of secretor and non-secretor phenotypes, which has not been investigated before in Kenya. Secondly, the study investigated the impact of red cell antigens expressed in mucosal surfaces (secretor status) on HIV infections, among a high-risk population of female sex workers. In addition the correlation between secretor status and HIV infection was correlated to CD4+ T cell counts on the basis, CD4+ T cell counts are the hallmark of disease progression. The final study objective was to screen for a *Secretor FUT2* gene null allele. This specific *se428* polymorphism is postulated to confer 1) a degree of protection against HIV infections 2) the non-secretor phenotype among individuals of African descent and 3) slower disease progression among HIV infected individuals.

To date, research centred on the role of HIV and red blood cells and corresponding antigens, has been minimal yet studies continue to demonstrate correlations between red cell antigens and susceptibility to infections. The infections associated with the variable expression of these antigens, are often mucosal, including Norwalk virus infections, Crohn's disease, *Vaginal candidiasis*, asthma and pheumonia. These studies suggest a person's secretor status; conferred by the *Fucosyltransferase 2* (*FUT2*) gene may be associated with varying disease susceptibilities. In some cases, the ABH secretors are more susceptible to infection, and in other instances the contrary is observed.

This raises a number of questions on what role blood group antigens play at the level of interactions with pathogenic organisms on mucosal surfaces, which are often the first portal of entry for microorganisms including HIV. Furthermore, pathogens are always evolving, however, they are also highly selective agents. Therefore for microorganisms that preferentially bind to carbohydrate moieties, host cell surface molecules such as the ABO blood group antigens can reveal patterns of selection. This study, therefore determined blood group phenotype frequencies, secretor and non-secretor phenotypes in the Kenyan population. This knowledge may inform disease risk and susceptibility patterns; and contribute towards the elucidation of the underlying 'protective' effect conferred by either the secretor and non-secretor status to infections including HIV which continues to exert a devastating toll within this setting.

Secretor status relates to the expression of ABH blood group antigens in the mucosa. In relation to this, HIV is primarily a mucosal infection. Therefore, curbing the HIV/AIDS pandemic would imperatively incorporate effective efforts and measures addressing the mucosal transmission of infection. This study therefore investigated the specific role of the ABH antigens in the genital mucosa in relation to HIV infections.

The female genital tract was selected as the focus of the study, because the FGT site accounts for the highest proportion of heterosexually transmitted infections. Furthermore, the Kenyan epidemic continues to present a feminized trend with females accounting for nearly twice the number of HIV infections within the region. Zeroing in, the modes of transmission study (2008) highlighted female sex workers and their clients' alone account for nearly 14% of new infections annually. Therefore, studies geared towards curbing the HIV epidemic, focusing on the female genital tract, among female sex workers, would

have a huge public health impact not only on the entire HIV epidemic in Kenya but across the sub Saharan Africa region.

Host genetics are also known to have an impact on the susceptibility and course of HIV infections. In relation to secretor status, and underlying genotypes, the null *FUT2* allele *se428*, has also been strongly associated with increased susceptibility to HIV infections and slow progression to disease. In addition, this is the most common SNP associated with the non-secretor phenotype in African populations is the non-functional *FUT2 se428*, which results in a stop codon and thereby production of a non-functional form of the enzyme alpha (1,2) fucosyltransferase. Therefore, as single nucleotide polymorphisms (SNPs) are the most widely used markers in studies to assess inter-individual genetic variations and among a people, this study profiled for this particular *FUT2* null allele, to better understand the impact of *FUT2* genetics on HIV infections and secondly determine the extent to which mutations at this site determine the phenotypic profiles seen in Kenya.

This study was designed to determine the baseline prevalence of the major ABO blood group and minor Duffy antigens in the Kenyan population. The study further profiled secretor and non-secretor prevalence rates; secretor status was determined at two comparative sites, oral and the female genital tract. The study focus was on the expression of blood group antigens on the female genital tract epithelium, the earliest point of HIV-host contact, and potential associations with the establishment of HIV infection. In addition to the above, the CD4+ T cells are known as the hallmark of disease progression markers. This study therefore sought to investigate the role of secretor phenotypes in the context of HIV infections, based on CD4+ T cell counts as a surrogate

marker of disease progression. This multi-faceted approach, allowed the comprehension of the impact of both red blood cells and T cells, on HIV infection rates.

Finally, to date, the molecular description of variation of *FUT2*, in the Kenyan population has not been undertaken. Thus the novelty of this work was the characterization of secretor and non-secretor genotypes based on an optimized next-generation sequencing protocol among study participants in Nairobi, Kenya. Genomic DNA samples were sequenced and the genotypes compared with the wild type *Se* allele. The focus was on the non-functional *se428* null-allele that has been associated with the non-secretor status in African populations.

Kenya is an ideal setting to conduct HIV research relevant to sub-Saharan Africa, where the challenges of the HIV pandemic are more acute than in other parts of the world. Sex workers have multiple sex partners and co-factors for HIV acquisition, placing them at increased risk of HIV infections. This was therefore a very relevant study population to address these research questions. If any protective associations are observed with blood antigen expression profiles, understanding of the mucosal mechanism of protection may provide insight into the development of new HIV preventive technologies.

1.2 Hypothesis

There are distinct correlations between *Secretor FUT2* gene polymorphisms and blood group antigen secretor status in Kenyan populations and blood group antigen non-secretors are less susceptible to HIV infections.

1.2.1 Sub Hypotheses

1. There are novel *Secretor FUT2* gene polymorphisms in Kenyan populations, and distinct correlations between *Secretor FUT2* gene polymorphisms and blood group antigen secretor status

2. There are correlations between blood group antigen expression profiles and HIV susceptibility; non-secretors are less susceptible to HIV infections

1.3 Objectives

1.3.1 Broad Objective

To investigate associations between mucosal blood group antigen expression profiles (secretor status), *Secretor FUT2* gene polymorphisms and susceptibility to HIV infection among female sex workers in Nairobi, Kenya.

1.3.2 Specific Objectives

1. To determine ABO, Duffy and Rhesus (D) blood group antigen frequencies among voluntary non-remunerated blood donors in Nairobi, Kenya.
2. To determine the frequency profiles of secretor and non-secretor phenotypes in the Kenyan population.
3. To investigate correlations between secretor status and HIV infection rates among a high-risk cohort of female sex workers in Nairobi, Kenya.
4. To investigate correlations between secretor status and HIV infection based on CD4+ T cell counts as a surrogate marker of disease progression.
5. To investigate the genetic variations and underlying polymorphisms among a population of female sex workers in Nairobi, Kenya; with a focus on the *FUT2 se*^{428} short nucleotide polymorphism.

2 LITERATURE REVIEW

2.1 The Human Immunodeficiency Virus

Human Immunodeficiency Virus (HIV) was discovered in 1984 by two independent labs as the causative agent of Acquired Immunodeficiency Syndrome (AIDS). AIDS is a clinical condition characterized by a decline of $CD4^+$ T cells, general immune dysfunction, and increased susceptibility to opportunistic infections and malignancies. Although there may be vague symptoms associated with initial viremia, the rate of onset of clinical disease following infection by HIV is variable; infected individuals may remain asymptomatic for a decade or more before the immunodeficiency becomes severe enough to cause illness (Palmisano and Vella, 2011; Haase, 2010; Hladik, 2008; Dieffenbach and Fauci, 2011).

2.1.1 Types, Groups and Subtypes of HIV

HIV is a member of the genus *Lentivirus* part of the family *Retroviridae* and is known to mainly infect CD4+ T cells, but also to a lesser extent monocytes, macrophages and dendritic cells. Once infected, the cascade of infections begins resulting in general immune dysfunction and immune-suppression (Haase, 2010; Hladik, 2008).

Types of HIV

There are two major types of HIV: HIV-1 and HIV-2. HIV-1 accounts for most infections globally, while HIV-2 is endemic in West Africa (Grez *et al.*, 1994). HIV-1 and HIV-2 are similar although they differ in molecular weight and the respective accessory genes. Furthermore although both types eventually result in progression to AIDS, HIV-2 is far

less pathogenic in comparison to HIV-1 (Rowland-Jones and Whittle, 2007).

Strains of HIV-1

The degree of diversity seen in HIV-1 is greater than that of any other virus observed (McBurney and Ross; 2008). Studies in the early 1990s established HIV-1 strains can be classified into four groups: the "major" group M, the "outlier" group O and two new groups, N and P (Charneau *et al.*, 1994; De Leys *et al.*, 1990; Gurtler *et al.*, 1994; Vanden Heasevelde *et al.*, 1994). Group O appears to be restricted to West and Central African regions. Group N was recently discovered in Cameroon and is extremely rare and a new strain closely relating to gorilla Simian Immunodeficiency Virus (SIV) was isolated from Cameroon and was designated group P (Simon *et al.*, 1998; Roques *et al.*, 2004; Plantier *et al.*, 2009; Ayouba *et al.*, 2001; Vallari *et al.*, 2011). Most infections belong to HIV-1 group M. Within group M there are known to be at least nine genetically distinct subtypes (or clades) of HIV-1. These are subtypes A, B, C, D, F, G, H, J, K and Circulating Recombinant Forms (CRFs). East Africa is predominated by subtypes A and D as well as derived recombinants. Subtype A is quite successful in Kenya; A and D in Uganda; A, D and C recombinants in Tanzania, while southern African countries are mainly predominated by subtype C (Taylor *et al.*, 2008; Shao and Williamson, 2012).

2.1.2 Structure of HIV-1

HIV-1 is a diploid virus and contains two molecules of positive sense, single-stranded RNA. The HIV-1 genome conforms to the standard retrovirus gene order 5'-gag-pol-env-3' in addition to regulatory genes. Moreover, the virion contains the enzyme Reverse

Transcriptase (RT), which is responsible for the production of one double stranded provirus DNA molecule from the HIV-1 RNA. This provirus DNA molecule is then integrated into the host genome during virion replication (Dimmock, Easton and Leppard, 2007; Crowley, 2009).

2.1.3 HIV-1 Structural Proteins

HIV-1 structural proteins are encoded by the *gag, pol* and *env* genes; *gag* encodes the virion core proteins, *pol* encodes the virion enzymes, and *env* encodes the viral envelope protein (Checkley *et al.*, 2011) as shown in Figure 2.1.

Virion Proteins

Gag: The *gag* gene encodes the Gag proteins form the matrix (MA, p17), CA (p24) and the p7 Nucleocaspid. The p17 matrix is located just below the viral membrane (envelope) forming the inner shell. Within the core, there is a cone shaped p24; p24 encloses the viral genomic RNA. The p7 Nucleocaspid interacts with the viral RNA inside the capsid (Dimmock, Easton and Leppard, 2007; Crowley, 2009; Checkley *et al.*, 2009).

Pol: The *pol* gene encodes viral proteins p17, p24 and p7 are generated by the action of viral aspartyl protease (PR/Pro, a product of *pol*) processing the HIV-1 p55 Gag precursor polyprotein. The Gag proteins on the p55 are located in the following order: p17-p24-p2-p7-p1-p6 (Dimmock, Easton and Leppard, 2007; Crowley, 2009). Polyproteins from *env* are cleaved by a host cellular protease. The nucleocaspid is also known as the gag caspid. Within this nucleocaspid are two identical RNA strands. These

RNA strands are closely associated with Pol, the RNA-dependent DNA polymerase also known as reverse transcriptase (p66, p51), Nucleocaspid (p7) proteins (p9 and p6) and MA (p17) is required for the incorporation of the Env proteins onto the mature virion (Dimmock, Easton and Leppard, 2007).

2.1.4 Envelope Proteins

Env: The *env* gene encodes the viral envelope glycoprotein (Env). This is the only viral protein found on the surface of the virion. It is synthesized as a precursor protein known as glycoprotein 160 (gp160), which is the precursor protein for gp120 and gp41. In the host cell, gp160 folds and trimerizes in the endoplasmic reticulum; obtaining ten disulfide bonds and approximately 30 N-linked glycans. This protein is then transported to the Golgi complex where a cellular protease cleaves gp160 into a soluble subunit, gp120 and a transmembrane unit, gp41. These subunits are non-covalently associated on the surface of infected cells and on HIV-1 virions (Dimmock, Easton and Leppard, 2007; Lever, 2007; Crowley, 2009).

2.1.5 Additional Proteins

There are a number of additional proteins closely associated with the core.

Vif: This gene codes for "virion infectivity factor", a protein that increases the infectivity of the HIV particle (Argyis *et al*., 2004). It is estimated that about 7 to 20 molecules of Vif are present per HIV virion

Vpr (Viral protein R for HIV-1) and Vpx (for HIV-2): These two exist as accessory viral gene products. The specific role and significance of these proteins has not yet been

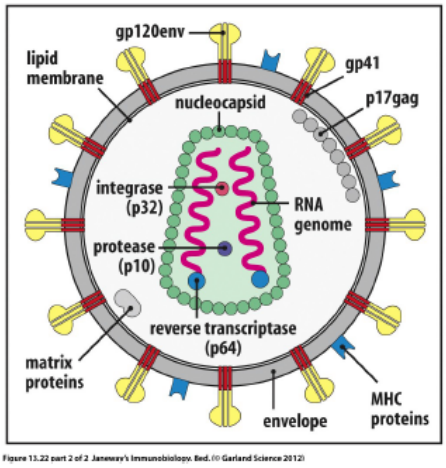
established to date but they are presumed to play a significant role in the early stages of HIV infections, accelerating the production of HIV proteins. (Ueno *et al.,* 2003; Dimmock, Easton and Leppard, 2007; Crowley, 2009).

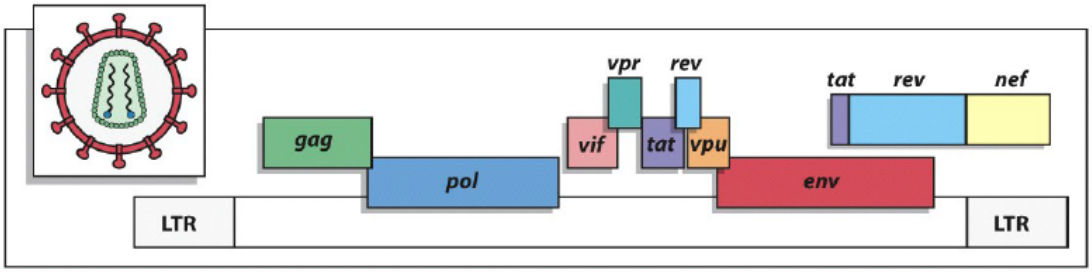
Vpu: The *vpu* gene encodes for "Viral protein U" helps with the assembly of new virus particles, and helps them to bud from the host cell.

Tat: Positive regulator of transcription, accelerates the production of more HIV virus. It's crucial to HIV, because HIV completely fails to replicate itself without it.

Rev: Regulator of viral expression, stimulates the production of HIV proteins, but suppresses the expression of HIV's regulatory genes. Rev also allows export of unspliced and partially spliced transcripts from the nucleus.

Nef: Negative regulation factor, this gene encodes a protein, which hangs around in the cytoplasm of the cell, and retards HIV replication.





Gene		Gene product/function
gag	Group-specific antigen	Core proteins and matrix proteins
pol	Polymerase	Reverse transcriptase, protease, and integrase enzymes
env	Envelope	Transmembrane glycoproteins. gp120 binds CD4 and CCR5; gp41 is required for virus fusion and internalization
tat	Transactivator	Positive regulator of transcription
rev	Regulator of viral expression	Allows export of unspliced and partially spliced transcripts from nucleus
vif	Viral infectivity	Affects particle infectivity
vpr	Viral protein R	Transport of DNA to nucleus. Augments virion production. Cell-cycle arrest
vpu	Viral protein U	Promotes intracellular degradation of CD4 and enhances release of virus from cell membrane
nef	Negative-regulation factor	Augments viral replication *in vivo* and *in vitro*. Decreases CD4, MHC class I and II expression

Figure 2.1: Structure of the Human Immunodeficiency Virus (Source: Janeway's Immunobiology, 2011)

2.2 Transmission of HIV Infections

2.2.1 Modes of Transmission

HIV is transmitted from an infected person via contact with infected body fluids including blood, blood products or organ/tissue transplants infected with the virus. The virus can also be transmitted by direct injection into the blood stream, or from an infected mother to a foetus or child during pregnancy, childbirth or breast feeding (Wood *et al.*, 2013). The primary mode of transmission, accounting for more than 80% of new infection is sexual transmission, both heterosexual and homosexual. However, people who inject drugs also account for the high-risk populations (Hladik *et al.*, 2007; UNAIDS, 2013).

2.2.2 Sexual Transmission of HIV

Sexual transmission of HIV is the primary route of transmission. This following exposure to either cell-free or cell-associated virions on genital mucosal surfaces (Hladik *et al*., 2007). There are a number of factors that favour the rapid dissemination of HIV including the absence of pre-existing immune pressures, which subsequently results in a burst of viremia often manifested during acute HIV infection. It is postulated that the initial infection is associated with an incompletely effective immune response to HIV. This results in the rapid increase in plasma viral numbers, and the subsequent resulting in establishment of a systemic infection (Moir *et al.* 2011).

Despite the great advancements in HIV research, the underlying events regarding the earliest events in HIV transmission following heterosexual transmission, in the female genital tract, have not been clearly elucidated. There has been insight, gleaned from *in*

vivo models based on Simian Immunodeficiency Virus (SIV) infections (Haase, 2005). There have also been epidemiological studies and *ex vivo* models mainly centered on understanding the initial target cells and the factors affecting transmission of HIV (Hladik and McElrath, 2008). Initial studies, based on SIV models by Zhang *et al.* (1999) identified $CD4^+$ T cells as the primary targets of productive viral replication, less than one-week post exposure. A few days following this initial event, the local propagation of the virions in the less abundant but more susceptible $CD4^+$ T cells takes place. However, based on SIV studies in macaques, following infection, virions rapidly migrate, probably via draining lymph nodes, to the gut-associated lymphoid tissue (GALT) (Li *et al.* 2005; Mattapallil *et al.* 2005). In the GALT the virus induces a massive depletion of $CD4^+$ T cells in the intestinal lamina propria. These events (i.e. rapid $CD4^+$ T cell depletion) are postulated to occur in humans as well.

Infection of target cells requires the interaction of the virion with CD4 as a receptor and a chemokine co-receptor (CCR5/CXCR4). During heterosexual mucosal transmission, the presence of $CD4^+$ T cells expressing CCR5, therefore, dictates the ability of HIV to establish a productive infection (Soeters *et al.*, 2013; Wheeler, 2013). The process whereby the virions penetrate the mucosal barriers to establish infection is still under research, both *in vivo* and using animal models. However, studies have demonstrated, based on *ex vivo* co-culture models, Langerhans cells (LCs), Dendritic cells (DCs) and/or Macrophages represent cell subsets present at the mucosal portal of entry (Jin *et al.*, 2014).

Interestingly, a study by Hladik *et al.* (2007), based on the early events of transmission in organ explants, demonstrated that HIV bound to and protected by Langerhans cells

(LCs), can in fact, cross the epithelial cell layers and further form complexes with the memory CD4+ T cells. In this case, forming the basis of productive viral replication. There are a number of additional factors, which have been proposed to influence HIV in establishment of infection. These include the proposed concept of virologic synapse, where in close proximity, additional virus-attachment molecules including integrins aid in HIV establishing a productive infection (Arthos *et al.* 2008). Piguet and Steinman (2007), support this line of thought, but based on the protective endosomal compartments within Langerhans Cells and Dendritic cells, which also help HIV in establishing a productive infection.

Findings based on epidemiological studies, have also contributed greatly into providing insight into the transmissibility of HIV. Gray *et al.* (2001) and Shattock *et al.* (2003) based estimates of rates of transmission on the frequency of coital acts in discordant couples, while findings by Hladik and McElrath (2008), Quinn (2007) and Tobian and Quinn (2009) examined host factors that may influence (increase or decrease) transmission. Moreover, studies have now found that circumcision does offer a degree of protection against HIV infection (Quinn, 2007; Moir *et al.* 2011). The latter findings are mainly based on the cell subsets, namely Langerhans Cells, Dendritic Cells, Macrophages and $CD4^+$ T cells, thought to increase the risk of infection by facilitating transmission (Quinn *et al.* 2007; Moir *et al.* 2011; Hladik and McElrath, 2008).

Additional underlying factors are the presence of underlying inflammation in the female genital tract, increasing the number of immune cells (HIV targets). Also, the virus breach may be further favoured by the presence of genital-ulcerative sexually transmitted infections/diseases (Tobian and Quinn, 2009).

2.3 Entry of HIV Into A Host Cell

Gp120 and gp41 mediate viral entry into the host target cells during infection (McDougal *et al.*, 1986; Lasky *et al.*, 1987). Gp120 proteins are assembled into trimers, termed the spike protein, at the distal part of the HIV-1 virion envelope. Each gp120 comprises five variable loops, which vary in sequence between HIV strains. Each virion has approximately 72 spikes. Gp120 binds to the cell surface receptor CD4, and the co-receptor CCR5 or CXCR4 on target cells. Gp41 is the trans-membrane subunit that plays an important role in anchoring gp120, and is important in the subsequent fusion of the virus to the host cell membrane (Clapham and McKnight, 2002; Crowley, 2009; Huang *et al.*, 2014).

During viral assembly these envelope proteins are transported to the host cell surface. Here part of the central and amino terminal portion of the transmembrane gp41 is expressed on the outside of the virion. This central region is the portion that binds to gp120 non-covalently; primarily at two hydrophobic regions. Gp41 also has a loop structure segment, which is important in membrane function. It is postulated that the long cytoplasmic tail of gp41 is required for HIV-1 envelope glycoprotein incorporation into the virions (Roche *et al.*, 2014; Qiu *et al.*, 2013). Gp120 is very important as it contains binding sites for the target cell receptors and the major antibody-neutralising domains. Although, the external gp41 domain has been reported to be contain neutralising antibody epitopes as well (Huang *et al.*, 2014). Interestingly, virion isolates have demonstrated the selective incorporation of specific lipid domains and host cell proteins from the host cell membrane during the viral budding process.

2.3.1 Reverse transcription and Integration

Following entry, the viral capsid undergoes uncoating, and release of the single stranded viral RNA into the cytoplasm of the host cell. Reverse transcription is initiated by the annealing of the 3'-terminus of cellular RNA primer to the complementary primer binding site of the viral RNA, forming a primer/template complex recognized by the viral reverse transcriptase. The viral RNA is then reverse transcribed into a single followed by a double-stranded cDNA by the viral reverse transcriptase in concert with RNAseH which degrades the original RNA molecule during the synthesis of the second strand of the DNA. The viral DNA is complexed with integrase, matrix, reverse transcriptase and Vpr to form the preintegration complex (PIC) that enters the cell nucleus through nuclear pores (Gotte *et al.*, 1999; Jonckheere *et al.*, 2000). Once inside the cell, the HIV enzyme called reverse transcriptase converts the viral RNA into DNA. This DNA is then transported to the cell's nucleus, where it is inserted into the host cell DNA by the HIV enzyme integrase. Once inserted, the HIV DNA is known as provirus. Infection of host cells starts the cascade of events that eventually result in immune dysfunction.

2.3.2 Transcription and translation

The integrated viral DNA directs the transcription of viral RNAs, which are transported into the host cell cytoplasm where translation of the viral proteins takes place. The newly synthesized viral proteins, together with two single- stranded copies of full-length (unspliced) viral RNA, assemble into a new generation of viral particles. Concomitant with release of virus particles from the infected cell, a third viral enzyme, protease (PR), cleaves the Gag and GagPol polyprotein precursors, triggering the conversion of the

immature particle to the mature virion. The virus replication cycle is now complete, and the mature virus particle can initiate a new cycle of infection (Freed *et al.*, 2006; Swanstrom *et al.*, 1997). This is summarized in Figure 2.2 below:

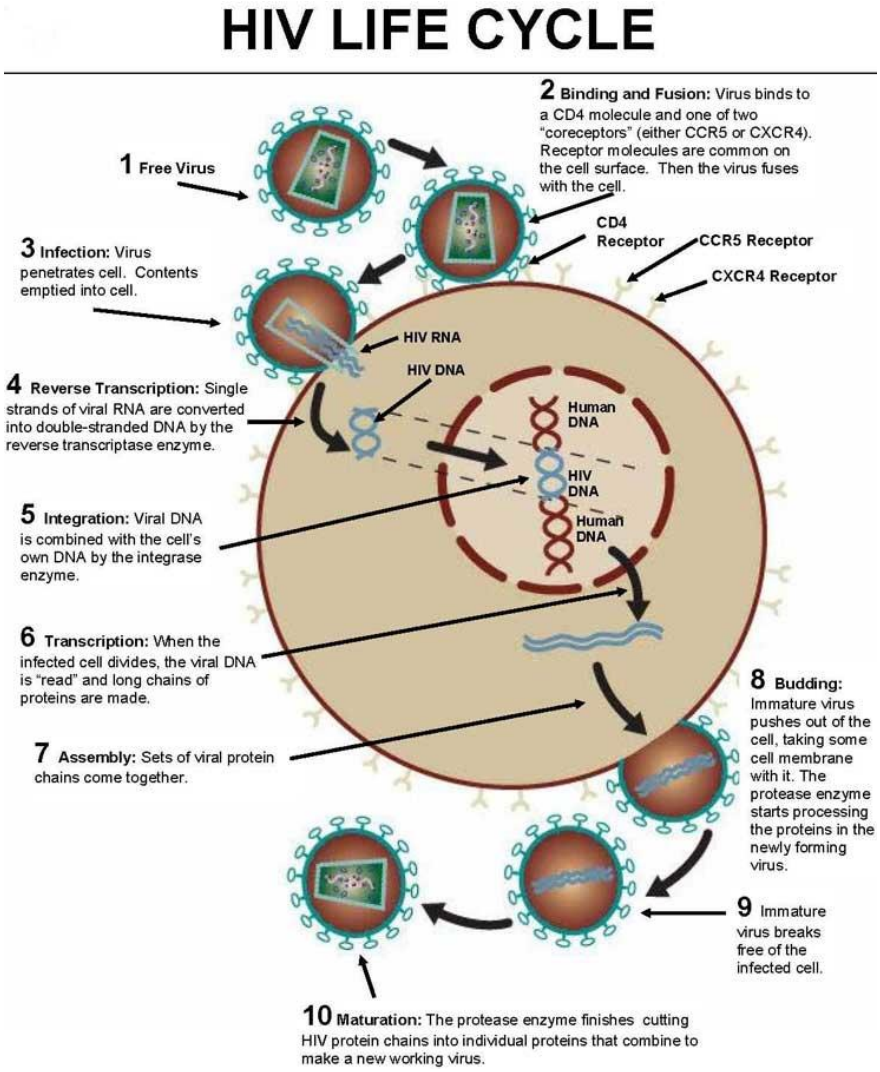


Figure 2.2: Infection of a Host Cell by The Human Immunodeficiency Virus (Janeway's Immunology)

2.4 Stages of HIV Infections

2.4.1 Acute HIV Infections

Following approximately 2 – 4 weeks post exposure, HIV infection is often characterized by experiencing an acute HIV syndrome defined by flu-like clinical manifestations associated with a high plasma viremia, fever and often lymphadenopathy (Gurunathan *et al.* 2009). Muscle pain (myalgias), skin rash, headaches, anorexia and diarrhea, may also present in patients, however, with a variable degree of severity in manifestation. During this phase of infection, there is an extremely aggressive replication of the virus, often favoured by the absence of immune pressures. In fact, the plasma viral loads can be as high as 10 million copies per millimeter (Moir *et al.* 2011; Little *et al.* 1999). In most cases, the dynamics of acute HIV infection (in the absence of Anti Retroviral Therapy), is characterized by a steady peak in plasma viremia, 3- 4 weeks post transmission. This is followed by a decline over the next few months, eventually reaching steady state, known as the viral set point (Little *et al.* 1999; Fiebig *et al.* 2003). The viral set point is particularly important, as a measure of the rate of disease progression in patients not treated with ART (Moir *et al.* 2011).

2.4.2 CD4+ T Cells During Acute HIV Infection

The hallmark of acute HIV infection is a dramatic depletion of $CD4^+$ T cells in the peripheral blood. The infection often results in the rapid infection of CD4+CCR5+ T cells in the draining lymph nodes. The virus then spreads quickly to other lymphoid tissues, preferentially infecting and significantly depleting memory CD4 T cells in multiple sites including the Gut-Associated Lymphoid Tissue (GALT), which are major

targets of HIV infection as high proportions of these cells express CCR5 and are activated. Studies based on SIV animal models, have associated GALT cell depletion, to the infection and direct killing, of the highly abundant, susceptible target cell population (in the GALT); often intensified by the extensive killing of bystander cell populations via apoptosis (Li *et al.* 2005; Mattapallil *et al.* 2005). Replication of HIV in the GALT, is then, followed by dissemination of the virus to the peripheral lymphoid tissues especially the lymph nodes. However, the depletion of $CD4^+$ T cells is less pronounced in the peripheral lymph nodes compared with the GALT (Brenchley *et al.* 2004). The virus is now able to establish lymphoid tissue viral reservoirs. During this time, the latent reservoir is also established. A pool of latently infected, resting CD4 T cells that carry integrated HIV-DNA that can produce replication-competent virus on activation.

2.4.3 Chronic HIV Infections

Massive immune activation and an accelerated cell turnover take place during chronic HIV infection (Ho *et al.*, 1995; Ford *et al.*, 2009). This apparent state of basal immune hyper-activation in the infected host is evidenced by increased expression of activation markers, such as CD38, HLA-DR and Ki67, of which CD38 is considered the most reliable surrogate marker for immune activation, disease progression to AIDS, and death (Lu *et al.*, 1997). In the gut, naïve and central memory T cells are supplied, but these cells are short-lived and only partially substitute for the CD4+ effector memory T cells depleted during the acute phase of infection (Grossman *et al.*, 2006). Immunological damage to the gastrointestinal tract leads to breaks in the mucosal barrier allowing translocation of microbial products, including bacterial lipopolysaccharide, into the

circulation. Bacterial translocation may therefore represent a crucial event in persistent immune activation, although it is probably not the only source of the microbial burden responsible for chronic immune activation. HIV itself may also be a central player in the process due to viral constituents, such as gp120 and *nef*, or viral nucleic acids produced during viral replication, subsequently resulting in activation of proinflammatory cytokines and type I interferon (IFN), including IFN-α and IFN-β (McMichael *et al.*, 2010; Boasso *et al.*, 2008)**.** The accelerated viral evolution at this stage, provided by an excessively high viral mutation rate and alteration in cellular tropism, results in progression from a pool of CCR5-trophic to dual trophic or dominantly CXCR4 trophic strains with increased virulence and broader target cell tropism. In addition, damage to lymphoid tissue results in thymic dysfunction, transforming growth factor-β-dependent fibrosis and alterations in lymphoid follicle architecture (Estes *et al.*, 2007). HIV infection also profoundly affects blood and tissue B cells by inducing early class switching in polyclonal B cells, massive B cell apoptosis, and loss of germinal centers in lymphoid tissue (He *et al.*, 2006).

2.5 Immunology of HIV Infections

Most people infected with HIV will mount an effective immune response to the virus during the first few months of infection. However, over time this response will prove ineffective. The responses come in two forms: cellular and humoral.

2.5.1 Cellular Immune Responses

Cellular immune responses refer to the activity of the CD4 and CD8 T cells, the latter known as cytotoxic lymphocytes. CD8 T cells act against HIV in two different ways during primary infection: by directly killing HIV-infected cells and by secreting anti-HIV molecules such as chemokines. While in most HIV infected people, CD8 T cell counts fall again after infection, some people continue to exhibit strong HIV-specific CD8 T cell responses, which control viral load. CD8 T cells play a crucial role in determining the speed of HIV disease progression (McMichael and Rowland-Jones, 2001). Most cytotoxic activity is carried out by CD8+ cells, which produce a cell antiviral factor, which can suppress viral gene transcription. CTL activity is directed against antigens on a variety of HIV proteins, including envelope glycoproteins, and internal structural proteins (p17, p24, p15), as well as against the products of regulatory genes such as nef and vif and pol (Collins *et al.*, 2014). CTLs are capable of lysing cells infected with HIV and have the advantage of a broad-based activity against isolates from different clades. CTL activity induced by HIV-2 can also lyse cells infected with HIV-1 (McMichael and Rowland-Jones, 2001; Sauce *et al.*, 2013).

CD4/CD8 cell ratio: To obtain a better picture of disease progression, it is advisable to measure the CD4/CD8 cell ratio. The normal ratio is about 1 to 1.5. If the disease is

progressing, CD4 will decline and CD8 will stay the same or increase and the CD4/CD8 ratio will be <1, indicating disease progression. Figure 2.3 below is a summary of the profile of CD4 and CD8 T cells during the course of HIV infection.

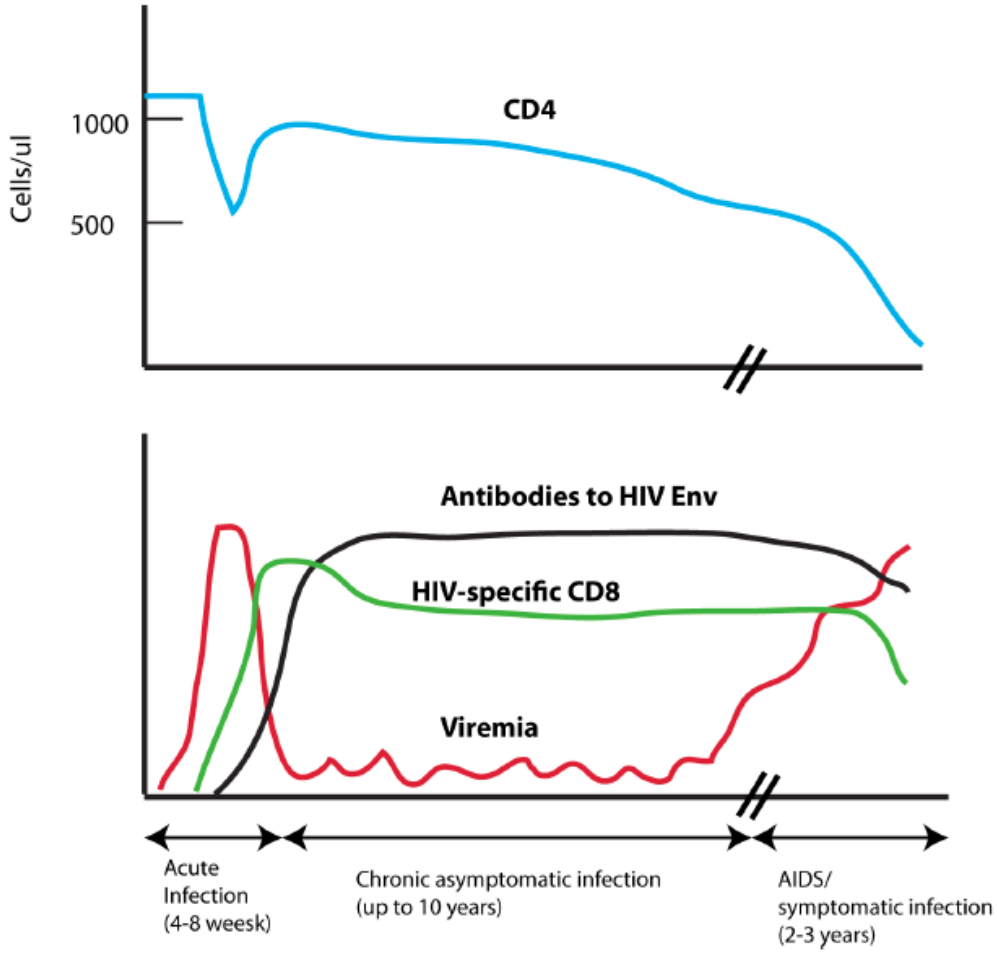


Figure 2.3: The phases of HIV/AIDS and the Magnitude of Immune Responses During the Course of Infections (Nature, 2006)

2.5.2 Humoral Responses

This refers to the production of antibodies against HIV. Antibodies to HIV can be detected soon after acute infection, often as early or a few days after exposure to the virus, but generally within 1 to 3 months. These antibodies can be found in the blood, on mucosal surfaces, and in various body fluids. IgG1 is a key player in host defense during

all stages of infection, and helps to combat the virus via antibody-dependent cellular cytotoxicity, complement-dependent cytotoxicity, and neutralizing and blocking responses (Mestecky *et al.*, 2014). Antibody-dependent cellular cytotoxicity is directed against sites on the viral envelope, especially against sites on the V3 loop and the extracellular domain of gp41. All the other antibody isotypes (IgM, IgA, IgG2, IgG4) vary in their levels throughout the course of infection (Gu *et al.*, 2014; Xu *et al.*, 2014). Neutralizing antibodies offer some protection by limiting viral replication during the early asymptomatic stages, but overall their titers tend to remain too low to clear the HIV infection. The low titer is probably a result of the fact that the virus's envelope epitopes have a highly dynamic configuration, which changes often, depending on the state of activation and binding to cellular receptors. The low titers also favour the emergence of resistant mutants during active replication (Xu *et al.*, 2014). Patients often produce titres of neutralizing bodies, often specific to the earlier virus isolates and therefore cannot neutralize "escape mutants" as effectively. Neutralizing antibodies bind to specific sites on the viral envelop complex, SU-TM. The V3 loop on SU is one of the chief targets of these immune molecules. Anti-V3 antibodies block co-receptor interactions that occur after the virus attached to the major CD4 receptor. Antibodies to this region tend to be very strain specific because of the high sequence variation in this segment (Gu *et al.*, 2014; Xu *et al.*, 2014).

2.5.3 The Host Immune System In Clearing HIV

There are currently numerous efforts all geared towards elucidating the mechanisms by which HIV fails to respond to immune pressures following exposure of a single virion to the host. Recent findings from studies designed and aimed to characterize the early founder transmitting virus in HIV-1 infection, have used single-genome amplification and mathematical models. Studies by Keele *et al.* (2008), Salazar-Gonzalez *et al.* (2009) and Goonetilleke *et al.* (2009), suggest that a single 'founder' virus (cell-free or cell-associated) is initially transmitted and that HIV begins to evolve or diverge from the founder virion, once a cellular immune response has been mounted against the pathogen; often several weeks post exposure.

It is postulated that transmission represents a genetic bottle-neck for the virus. There are certain traits from the founder virus, especially those of the viral envelope, which are preferentially selected. The initial founder virions are characterized by an increased sensitivity to neutralization and a scarcity of glycosylation sites on the virion envelope gp120 (Derdeyn *et al.* 2004). However, the immune response elicited, probably 12 weeks post exposure, characterized by the increase in production of neutralizing antibodies, is considered '*too little – too late*'. This is based on the fact that the initial virions have been quickly replaced in succession by neutralization-resistant variants (Karlsson Hedestam *et al.* 2008).

One of the key reasons, HIV is thought to favorably escape the host immune system, is the subsequent viral population following infection; characterized by virions that have highly glycosylated envelopes. This is the so-called '*glycosylation shield*', which prevents binding of neutralizing antibodies and promotes viral persistence (Wei *et al.*

2003). In line with this, a recent review by McMicheal *et al.* (2010), highlights that early HIV-specific $CD8^+$ T cell responses contribute to the decline in HIV plasma viral loads during the acute phase of infection. To counter this, there is a rapid evolution of the initially homogenous virus, which begins at the peak of the early HIV-specific $CD8^+$ T cell responses. This is characterized by mutations in epitopes recognized by $CD8^+$ T cells (Salazar-Gonzalez *et al.* 2009; Goonetilleke *et al.* 2009). It is important to note that this virion evolution takes place very early in the course of HIV infection, often before the elucidation of neutralizing antibodies. The continuation of this throughout disease progression contributes to infection (viral) persistence. Finally, Li *et al.* (2009), demonstrate evidence from animal models; additional factors such as location and magnitude of the immune response elicited are also likely to contribute to disease outcome. This places a key emphasis on the importance of HIV-specific T cell responses.

2.5.4 Host and Viral Factors in the Course of HIV Infection

There has been great insight on the influence, and role of host and viral factors on the course of HIV infections and progression to disease. Interestingly, questions remain unanswered about certain issue, including the HIV-resistant individuals – who maintain low to undetectable HIV plasma viremia and stable $CD4^+$ T cell counts, in the absence of ART. These are known as Long-Term Non Progressors (LNTPs). Within this group, there are individuals who maintain viral burdens below the limits of detection and are known as Elite Controllers. Majority of the studies on LNTPs and Elite Controllers focus on cellular restriction factors and viral fitness, in addition to host genetics and its impact on the ability to control virion replication, HIV infection, through innate, cellular and humoral immunity (Moir *et al.* 2011).

2.6 Vaccine Development

The development of an AIDS vaccine is affected by the range of virus subtypes as well as by the wide variety of human populations who need protection and who differ, for example, in their genetic make-up and their routes of exposure to HIV. In particular, the occurrence of super infection indicates that an immune response triggered by a vaccine to prevent infection by one strain of HIV may not protect against all other strains. The increasing variety of sub-types found within countries suggests that the effectiveness of a vaccine is likely to vary between populations, unless an innovative method is developed which guards against the circulating recombinant forms.

Inevitably, different types of candidate vaccines will have to be tested against various virion strains in multiple vaccine trials, conducted in both high-income and developing countries. To-date there have been three major vaccine efficacy trials. The AIDSVAX glycoprotein (gp)120 vaccine stimulated the production of non-neutralizing antibodies to the virus envelope proteins but failed to protect vaccinated individuals from infection (Pitisuttithum *et al.*, 2006). The STEP vaccine, comprised of three recombinant attenuated adenovirus serotype 5 viruses expressing HIV-1 *Gag, Pol* and *Nef*, stimulated $CD8^+$ T cell responses to the viral proteins but again showed no protective effect (McElrath *et al.*, 2008; Buchbinder *et al.*, 2008). Similar virus vector-based vaccines have been shown to stimulate simian immunodeficiency virus (SIV)-specific $CD8^+$ T cell responses in rhesus macaques, and an Adenovirus serotype 5 vector expressing Gag protected against challenge with a chimeric SIV–HIV (SHIV89.6p) virus but was not protective against challenge with the more natural SIVmac239 (Casimiro *et al.*, 2005).

Finally a third trial the RV144 study tested the safety and efficacy of a prime-boost

regimen comprising an ALVAC-HIV (a canary pox vector expression HIV-1 *env/gag/pro* prime and AIDSVAX-gp120 B/E (recombinant gp 120) boost in heterosexual individuals at various levels of risk of HIV infection (Rerks-Ngarm et al., 2009). These vaccines were designed to induce both antibody and cell-mediated immune responses and to complement each other in order to maximize protection. Hence, with 74 HIV infections among the placebo recipients compared to only 51 in the vaccines, RV144 achieved 31.2% efficacy and although there was no effect on viral load following infection, it so far remains the most encouraging vaccine study to date. Thus although highly elusive, the unsatisfactory outcomes of three large trials, mean that now more than ever, the world is more desperate for a safe and an effective vaccine to prevent HIV infection and/or control progression to AIDS.

Most recently, the focus has now moved to mucosal immunology. Data from numerous studies, point to the imminent possibility of a vaccine that can stimulate the greatly desired protective mucosal and systemic immune responses. It might be that carefully selected combination of immunogens, adjuvants, delivery vectors, and immunization routes may possibly yield an HIV-1 vaccine that induces optimal activation of the innate immune system, and elicit protective antibody and T cell responses in both the mucosal and systemic compartments. Mucosal immunization holds promise as the ultimate modality to ensure sustained levels of potent antibody and cellular immune responses at the genital mucosa, where they are required to arrest initial breakthrough infections. Moreover, since systemic responses do not accurately represent local immunity at the genitorectal mucosa, comprehensive immune-functional and phenotypic characterization of the mucosal anti HIV-1 immune response that correlates with *in vivo* virus inhibition,

together with the mechanisms involved, will be crucial to the design of an efficacious vaccine for HIV-1. These will include accurate quantification of threshold titres of the mucosal antibody and T cell responses that would be sufficient to prevent infection.

2.7 Epidemiology of HIV/AIDS

2.7.1 Global Burden of HIV

UNAIDS (2013) estimates there are currently over 30 million people are living with HIV/AIDS. The annual number of new infections in 2012 was 2.3 million a decline by 33% from 3.4 million (2001); and there was a witnessed decrease in AIDS related deaths to 1.6 million in 2012 from the 2.3 million witnessed in 2005 as summarized in Table 2.1 below:

	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
People living with HIV	30.0 million [27.2-33.1 million]	31.0 million [28.2-34.1 million]	31.7 million [28.9-34.8 million]	32.2 million [29.4-35.3 million]	32.5 million [29.7-35.6 million]	32.8 million [30.1-36.0 million]	33.2 million [30.4-36.3 million]	33.5 million [30.7-36.7 million]	34.0 million [31.1-37.1 million]	34.4 million [31.5-37.7 million]	34.9 million [31.9-38.3 million]	35.3 million [32.2-38.8 million]
New HIV Infections (Total)	3.4 million [3.1-3.7 million]	3.3 million [3.0-3.6 million]	3.1 million [2.9-3.5 million]	3.0 million [2.7-3.3 million]	2.9 million [2.6-3.2 million]	2.8 million [2.5-3.2 million]	2.7 million [2.4-3.1 million]	2.6 million [2.3-3.0 million]	2.6 million [2.2-3.0 million]	2.5 million [2.2-2.9 million]	2.5 million [2.1-2.9 million]	2.3 million [1.9-2.7 million]
New HIV infections (adults)	2.8 million [2.6-3.1 million]	2.7 million [2.5-3.0 million]	2.6 million [2.3-2.9 million]	2.4 million [2.2-2.7 million]	2.3 million [2.1-2.6 million]	2.3 million [2.0-2.6 million]	2.2 million [2.0-2.5 million]	2.2 million [1.9-2.5 million]	2.2 million [1.9-2.5 million]	2.2 million [1.9-2.5 million]	2.2 million [1.8-2.5 million]	2.0 million [1.7-2.4 million]
New infections (children)	550 000 [500 000–620 000]	560 000 [510 000–630 000]	560 000 [520 000–630 000]	550 000 [510 000–620 000]	540 000 [490 000–610 000]	520 000 [470 000–580 000]	480 000 [440 000–550 000]	450 000 [410 000–520 000]	400 000 [360 000–470 000]	360 000 [330 000–420 000]	310 000 [280 000–370 000]	260 000 [230 000–320 000]
AIDS-related deaths	1.9 million [1.7-2.2 million]	2.1 million [1.9-2.4 million]	2.2 million [2.0-2.5 million]	2.3 million [2.1-2.6 million]	2.3 million [2.1-2.6 million]	2.3 million [2.0-2.6 million]	2.2 million [1.9-2.5 million]	2.1 million [1.8-2.4 million]	2.0 million [1.7-2.3 million]	1.9 million [1.7-2.2 million]	1.8 million [1.6–2.1 million]	1.6 million [1.4-1.9 million]
People accessing treatment					1.3 million	2.0 million	2.9 million	4.1 million	5.3 million	6.6 million	8.1 million	9.7 million
Resources		US$ 3.8 billion	US$ 4.6 billion	US$ 5.7 billion	US$ 7.4 billion	US$ 8.8 billion	US$ 10.5 billion	US$ 14.6 billion	US$ 15.5 billion	US$ 15.6 billion	US$ 17.1 billion	US$ 18.9 billion

Table 2.1: Trend of HIV Infections, and AIDS related Deaths by Region (Source: UNAIDS 2013 Fact sheet
http://www.unaids.org/en/resources/campaigns/globalreport2013/factsheet/,

Accessed 23 April 2014 8.24pm)

2.7.2 The HIV Epidemic in Sub-Saharan Africa

On the global front, sub Saharan Africa (SSA) is the hardest hit, bearing an estimated two-thirds of the burden of the HIV epidemic with 25 million people living with HIV (Table 2) However, over the past decade there have been great leaps in the field of prevention within the region. There was a notable 40% decline from the 2.6 million (2001) to 1.6 million (2012). The AIDS related deaths also reduced by 22% from 1.5 million (2001) to 1.2 million (2012) owing to the life-prolonging effects of antiretroviral therapy (UNAIDS, 2013: AIDS by the Numbers).

Overall, HIV prevalence is on the decline in most countries but the effect of the diminishing HIV rates - are not strong or widespread enough to diminish the epidemics' overall impact. In Africa, countries in the Southern parts have the highest prevalence rates including Swaziland, Botswana, Lesotho, South Africa and Zimbabwe with country prevalence rates of 26.5%, 23.0%, 23.1%, 17.9% and 14.7% respectively while lower and more generalized epidemics are seen in Eastern Africa including Kenya (5.6%), Tanzania (5.1%) and Uganda (7.2%), Rwanda (2.9%) and the neighboring South Sudan (2.9%). Furthermore, the epidemic has a feminized trend, as the HIV prevalence rates among young women are twice as high as those of the young men.

In the field of prevention, there has been great scale up of antiretroviral therapy within SSA to nearly 60%, increment in biomedical prevention efforts including voluntary male circumcision with 3.2 million African men circumcised in 2012 (UNAIDS 2013, AIDS by the Numbers). Furthermore, although sexual transmission is the primary driver of HIV infections in the region, the trends in reduction via sexual transmission have generally been favourable. Changes in sexual behavior including delayed sexual debut, and the

adoption of safer sexual behaviours including increase in condom use, reduction in multiple partners have ultimately altered the natural dynamics of the epidemic (UNAIDS, 2013). The impact of HIV in SSA as compared to other global regions is outlined in Table 2.2 below.

Region	People living with HIV 2012		New HIV infections 2012		AIDS-related deaths 2012 (total)
	total	children	total	children	
Sub-Saharan Africa	**25.0 million** [23.5 million–26.6 million]	**2.9 million** [2.7 million – 3.3 million]	**1.6 million** [1.4 million – 1.8 million]	**230 000** [200 000-280 000]	**1.2 million** [1.1 million-1.3 million]
South and South-East Asia	**3.9 million** [2.9 million-5.2 million]	**200 000** [170 000-270 000]	**270 000** [160 000-440 000]	**21 000** [16 000-32 000]	**220 000** [150 000 – 310 000]
East Asia	**880 000** [650 000-1.2 million]	**8 200** [5 800-11 000]	**81 000** [34 000-160 000]	**1 500** [<1 000-3 300]	**41 000** [25 000-64 000]
Latin America	**1.5 million** [1.2 million-1.9 million]	**40 000** [32 000-52 000]	**86 000** [57 000-150 000]	**2 100** [<1 000-4 600]	**52 000** [35 000-75 000]
Western and Central Europe	**860 000** [800 000-930 000	**1 600** [<1 300-2 000]	**29 000** [25 000-35 000	**<200** [<100-<200]	**7 600** [6 900-8 300]
North America	**1.3 million** [980 000-1.9 million]	**4 500** [4 000-5 800]	**48 000** [15 000-100 000]	**<200** [<200-<500]	**20 000** [16 000-27 000]
Eastern Europe and Central Asia	**1.3 million** 1.0 million-1.7 million	**19 000** [16 000-24 000]	**130 000** [89 000-190 000]	**<1 000** [<500-1 200]	**91 000** [66 000-120 000]
Caribbean	**250 000** [220 000-280 000]	**16 000** [14 000-19 000]	**12 000** [9 400-14 000]	**<500** [<500-<1 000]	**11 000** [9 400-14 000]
Middle East and North Africa	**260 000** [200 000-380 000]	**20 000** [14 000-31 000]	**32 000** [22 000-47 000]	**3 000** [2 000-4 600]	**17 000** [12 000-26 000]
Oceania	**51 000** [43 000-59 000]	**3 100** [2 400-4 100]	**2 100** [1 500-2 700]	**<500** [<200-<500]	**1 200** [<1 000-1 800]
Global	**35.3 million** [32.2 million-38.8 million]	**3.3 million** [3.0 million-3.7 million]	**2.3 million** [1.9 million-2.7 million]	**260 000** [230 000-320 000]	**1.6 million** [1.4 million-1.9 million]

Table 2.2: Global burden of HIV by region (Source: UNAIDS 2013 Fact sheet
http://www.unaids.org/en/resources/campaigns/globalreport2013/factsheet/, Accessed 23 April 2014 8.24pm)

2.8 Burden of HIV In Kenya

Kenya faces a generalized HIV/AIDS epidemic, however as it is globally, the overall HIV/AIDS prevalence within the country is on the decline among the general population aged 15-49 years declined from 10% (Sentinel Surveillance) in 1998 among the general population to 5.6% in 2012 (KAIS 2012) with women displaying a higher prevalence at 6.9% versus 4.4% for men (KAIS 2012). The reduction in the prevalence of AIDS-related deaths has also decreased, which has been primarily attributed to an increase in access to Antiretroviral Therapy (ART), promising increases in safe behaviour such as increases in condom use, delay in sexual debut, and a reduction in the number of sexual partners. There are a number of factors that have a clear association with the prevalence of HIV/AIDS among the Kenyan population, detailed below:

Age and Sex

HIV prevalence rates increase with age and the trend is feminized with more females being infected with HIV as shown in the figures 2.4 and 2.5 below:

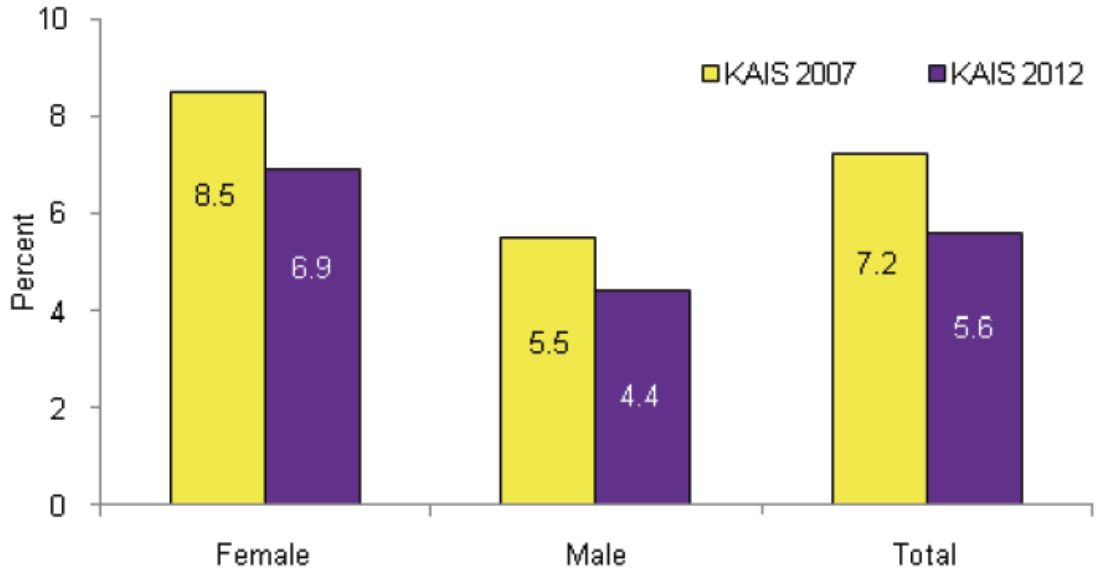


Figure2.4: HIV prevalence stratified by sex, in 2007 and 2012; the incidence of HIV is higher among the females (KAIS, 2012)

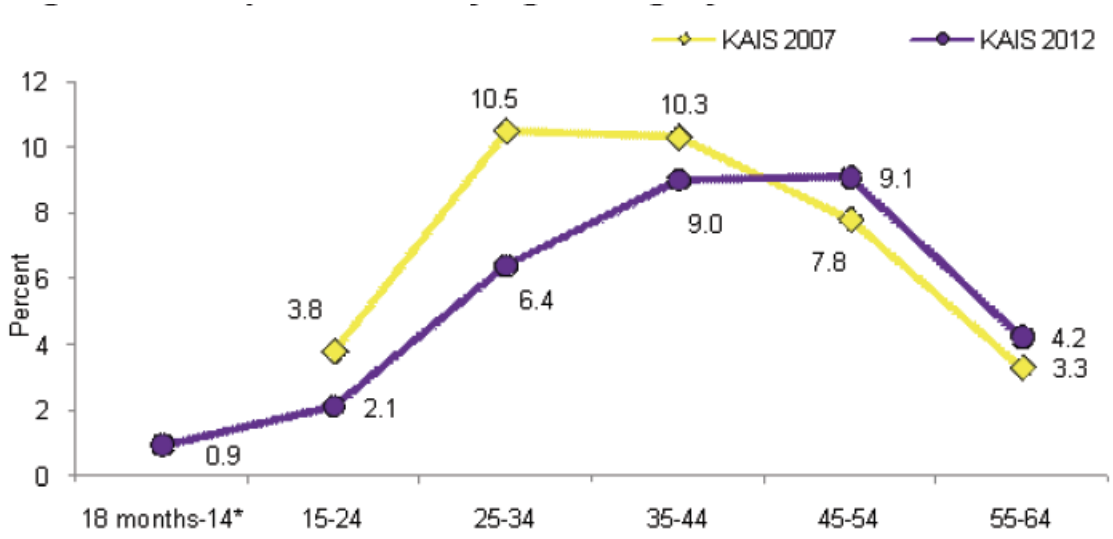


Figure 2.5: HIV prevalence stratified by age in 2007 and 2012; the incidence of HIV increases with age (KAIS, 2012)

Residence and Region

Kenya displays a mixed HIV epidemic, with prevalence varying widely across regions, and between urban and rural settings (KAIS, 2012). In the rural areas, HIV prevalence among adults aged 15-64 years was estimated at 5.1%, compared to 6.5% among those from the urban setting. However, because the vast majority (65%) of Kenya's population is based in rural regions, the absolute number of HIV infections is higher in rural compared to urban areas (1 million versus 0.4 million adults, respectively). The gender imbalance in HIV prevalence is reflected in both urban and rural areas, with urban-based women displaying a higher HIV prevalence compared to those in the rural setting (8.0% versus 6.2%, respectively). For the men, HIV prevalence is, urban against rural, 5.1% versus 3.9% respectively. HIV prevalence also varies substantially between regions; in some cases, HIV prevalence can differ by as much as 15-fold between regions. Figure 2.6 highlights the regional prevalence variation across Kenya.

2.8.1 HIV Prevalence across Regions in Kenya

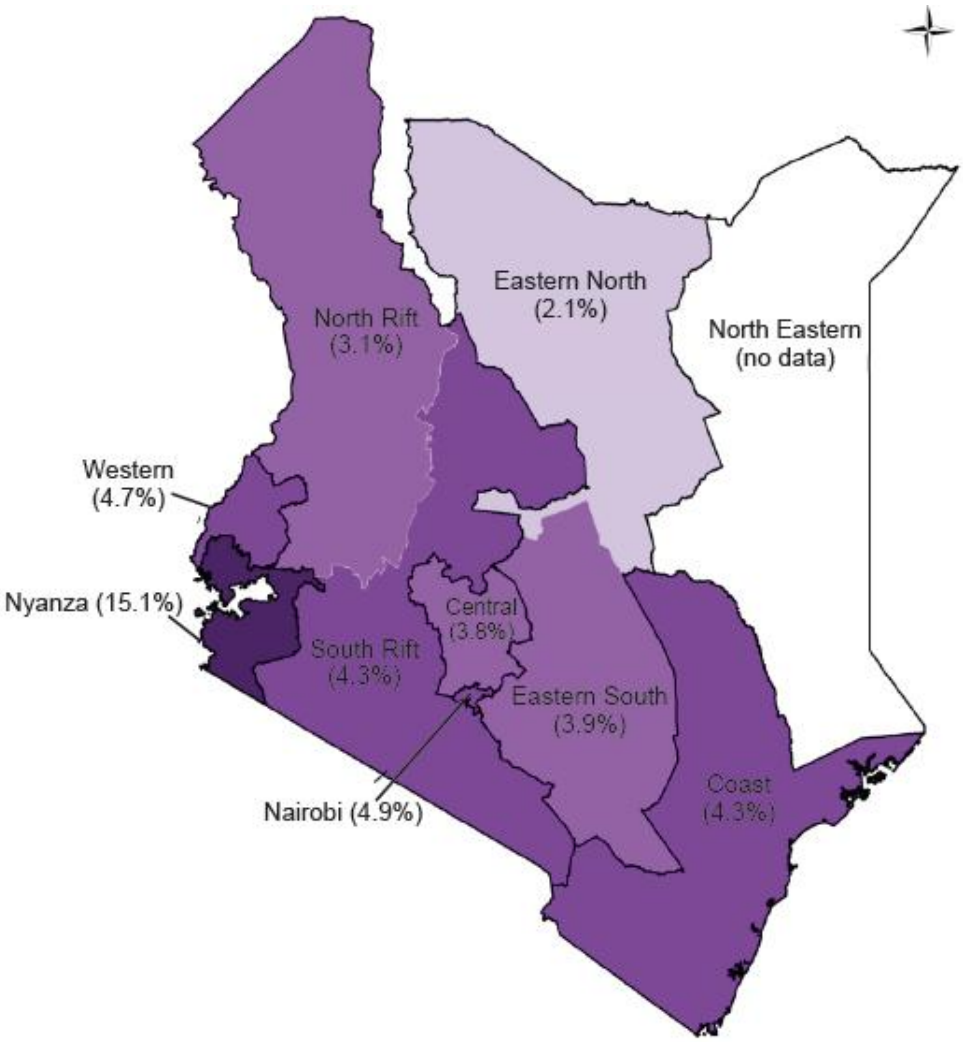


Figure 2.6: HIV across regions in Kenya (KAIS, 2012)

Marital Status

The Kenya AIDS Indicator Survey (2012), reports the highest HIV prevalence observed is among widowed respondents (20%), and lowest among those who are single and never married had the lowest (1.8%). The prevalence among those married or cohabiting was approximately 5.3% while among those who were not in any marriage unions, women had a prevalence of 2.7% while the men had a prevalence of 1.3%.

Most At Risk Populations in Kenya (MARPS)

In STI epidemiology, a small proportion of a given population, due to behaviours that lead to exposure, experience disproportionately high numbers of infections. For the current HIV epidemic, these include female sex workers (FSWs), their clients, men who have sex with men (MSMs) and people who inject drugs (PUDs). Currently, surveillance of these risk groups is weak, and there is a need to enhance the current strategies to determine more accurate prevalence data in these important populations. The MoT Spectrum model of 2008 highlights that the current HIV/AIDS epidemic crisis is being driven largely by the MARPs in Kenya; sex workers were estimated to contribute 14% of new infections, MSM and prison populations, and 15%, PUDs.

2.8.2 HIV Infections among High Risk Populations

Focus on Sub Saharan Africa

The trends in reduction of HIV infections via sexual transmission have generally been favourable in SSA. However, there has been an increase in transmission among high-risk populations including commercial sex workers, and efforts to reduce transmission among these key groups remain insufficient. This subpopulation has specific behavioural and biological risk factors including multiple sexual partners, numerous exposure episodes, incidence of ulcerative diseases and sexually transmitted infections which all have an impact on the concentrated HIV epidemics among them and thereby translating into the general population (Baral and Beyrer *et al*., 2012).

Zeroing in: globally female sex workers are 13.5 times more likely to be living with HIV than other women (Baral *et al*., 2012). This effect is more pronounced in SSA. In West African countries, sex workers account for 10-32% of new infections, Uganda, Swaziland and Zambia, 7-11%, Kenya 14% of new infections are attributed to sex workers and their clients (MoT, 2008).

The HIV prevalence among this high-risk population is also alarmingly high; pooled data found rates of HIV among FSWs is 36.9% (Kerrigan *et al*., 2010). In Eastern and Southern Africa (data from 8 countries) the HIV prevalence rate was 22% and in Western and Central Africa (17 countries) a rate of 17% while the rates were less than 5% in all other regions. A limitation of these findings is that they are based on urban population surveys and therefore are not a true reflection of the national prevalence rates among these high-risk populations.

There has been an overall increase in condom use among these women but the inadequate

financing on HIV prevention programmes focused on FSWs both from local and international governments remains low except in southern Africa where domestic spending on HIV prevention services for sex workers outweighs international contributions. Sex workers are a sub population at a disproportionate risk of acquiring HIV, compounded by social stigma and the legal disadvantages. It is therefore imperative for host country governments to translate the recognition of HIV prevention needs among sex workers, their clients and partners into scaled-up evidence-and rights-based programmes.

2.8.3 Focus on The Pumwani Majengo Sex Worker Cohort, Nairobi, Kenya

The *Pumwani cohort* was started in 1983. As of 2013, more than 60,000 contacts had been made, with more than 20,000 HIV tests performed. Approximately 9,000 volunteers were receiving care, with 4,000 on antiretroviral treatment; the HIV prevalence rate among this cohort is currently at 22%. Outreach activities continue with this population. Sex workers are contacted in the field and encouraged to visit clinics, enroll and access services. The volunteers are primarily recruited from the slum area of Pumwani, Nairobi, Kenya.

This cohort is unique in Africa for the duration and depth of follow-up of such mobile, high-risk women. Much work has been done with this cohort. A key finding early on was the phenomenon of a subset of these women who are highly exposed to HIV through their sex work but remain uninfected with the virus – these women are considered to be resistant to HIV infection. They are referred to as the HESN – Highly Exposed Seronegative. A second unique set of women, get infected with the virus, but maintain

low to undetectable HIV plasma viremia and stable $CD4^+$ T cell counts, in the absence of ART. These are known as Long-Term Non Progressors (LTNPs). Within this group, there are individuals who maintain viral burdens below the limits of detection and are known as Elite Controllers. Majority of the studies on LTNPs and elite controllers focus on cellular restriction factors and viral fitness, in addition to host genetics and its impact on the ability to control virion replication, HIV infection, through innate, cellular and humoral immunity.

2.9 Human Blood Group Systems

2.9.1 An Overview

The discovery of the human blood groups in the 1900s by Landsteiner led to the understanding of blood group antigens as patterns of inherited serological reactions (Landsteiner, 1900). Today, there are more than 20 distinct blood group systems, comprising approximately 400 antigens (ISBT, 2014). Of these, the ABO and Rhesus blood groups systems remain the most important clinically, in transfusion and transplantation medicine. These well-defined markers are also the most genetically polymorphic of all human blood group antigens to-date (Avent and Reid, 2000; Giri *et al.*, 2011; Segurel *et al.*, 2012).

The ABO antigens comprising A, B and H are carbohydrate moieties expressed on red blood cells depending on the activity and specificity of enzyme glycosyltransferases encoded by the *Fucosytransferase 1* (*FUT1*) gene (Yamamoto *et al.*, 1990a; Yamamoto *et al.*, 1990b). They catalyze the transfer of N-acetyl-D-galactosamine or D-galactose or both to the non-reducing ends of suitable oligosaccharide chains found on red cell membrane glycoproteins and glycolipids (Morgan and Watkins, 1953). This results in the expression of blood group A and B antigens respectively while blood group O results from inactivity of the glycosyltransferase gene that generates A, B or both antigens.

The distribution of blood group phenotypes differs among populations globally and has been clearly documented, however, the principle underlying the major distribution difference among populations is not fully understood (Rahman *et al.*, 2011; Lialiaris *et al.*, 2011; Hamed *et al.*, 2012; Temitayo and Timothy, 2013; Mwangi, 1999). Furthermore, there is now extensive knowledge relating to the ABH glycans and their

role in immunohaematology. The blood group O phenotype has been associated with a higher risk of peptic ulcers (Edgren *et al.*, 2010); Blood group A with increased *Plasmodium falciparum* malaria disease severity (Fry *et al.*, 2008); Blood group B, associated with Chagas disease, dengue fever and Human T-lymphotropic Virus-1 (HTLV-1) (Teixeira, *et al.*, 1987; Kalayanarooj *et al.*, 2007; Ayatollahi *et al.*, 2008) and type AB individuals have been shown to have an increased incidence of pancreatic cancer (Wolpin *et al.*, 2010).

In addition to the major ABO blood group system, the Rhesus blood group system comprises the second most immunologically important blood group antigens (Flegel, 2011). Dating back to the early 1940s, the Rhesus (Rh) antigens play a significant role in transfusion medicine (Levine *et al.*, 1941; Nance and Lomas-Francis, 2013). Rh antigens are protein motifs with the ability to mount potent alloimmune reactions in transfusion, pregnancy - haemolytic disease of the newborn (Avent and Reed, 2000; Flegel, 2011), and play a key role in sickle cell disease (Noizat-Pirenne, 2012). However unlike the ABO blood group system, anti-Rhesus D antibodies develop only when a Rhesus D negative person is transfused with Rhesus D positive blood (Crowther *et al.*, 2013). Once produced, these antibodies persist in blood for years and can produce serious reactions during a second transfusion, known as Rhesus incompatibility. In pregnant mothers, if the mother is Rhesus D negative and the child is Rhesus D positive, this results in haemolytic disease of the newborn; severity may worsen in subsequent pregnancies (Chown, 1954; Chilcott *et al.*, 2002). Similar to ABO frequency profiles, Rhesus D antigen distribution is variable globally. There is a notable low prevalence of the Rhesus D negative phenotypes among individuals of African descent (Pennap *et al.*, 2011; Erbahor *et al.*,

2010).

An additional blood group system is the minor Duffy blood group system comprising the Duffy blood group antigens. The Duffy antigen also known as the Duffy Antigen Receptor for Chemokines (DARC) is a high-affinity receptor for pro-inflammatory chemokines such as Interleukin-8 (IL-8), macrophage inflammatory protein-1 (MIP) and regulated and normal T cell expressed and secreted (RANTES) (Neote *et al.,* 1993). DARC expressed on erythrocytes is also known to be a receptor for malarial parasites *Plasmodium vivax* and *Plasmodium knowlesi* (Pogo and Chaudhuri, 2000). Furthermore, DARC has been demonstrated as a receptor for the Human Immunodeficiency Virus (HIV-1), as well as molecules such as tetraspanin CD82 (Bandyopadhyay *et al.,* 2006). Interestingly, the Duffy antigen-negative phenotype, Fy (a^- b^-), has a high prevalence in individuals of African ancestry, while remaining exceedingly rare among Caucasians (Howes *et al.*, 2011). This negative phenotype is characterized by lack of DARC expression on red blood cells and is due to a homozygous point mutation on the gene, in the erythroid specific promoter (Pogo and Chaudhuri, 2000; Daniels *et al.*, 2005).

Frequency studies of the major and minor blood group phenotypes including ABO, Rhesus and Duffy in families and across populations globally are therefore, multipurpose. The importance of these well-defined antigens is increasingly being sought in their relation to disease, forensic pathology, population genetics, and anthropological studies.

This study was therefore designed as a baseline profile of the ABO, Rhesus (D) and Duffy (Fya, Fyb) blood group antigens among a low risk population of voluntary blood donors, in Nairobi, Kenya.

2.10 The ABO Blood Group System

The ABO system was first described in the early 1900s, based on serological differences in blood between individuals. Interestingly, the antigens comprising this blood group system were among the first human genetic markers identified (Storry and Olsson, 2009; Koda *et al*., 2001; Ferrer-Admetlla *et al*., 2009). This major blood group system consists of four blood types: A, B, AB and O (Daniels *et al*., 2004). A, B and O were initially described by Landsteiner *et al.*, (1900), while the DesCasterllo and Sturli (1902) discovered the fourth blood group, AB. These antigens are genetic markers inherited as Mendelian characteristics in a codominant autosomal fashion.

The variation of these phenotypes among populations has been clearly documented. There is a high frequency of the A phenotype mainly in Northern and Central Europe. The B phenotype is more frequent in Asia, while the O phenotype is most common globally, with a distinct increase in Africa and Australia (Storry and Olsson, 2009). The principle underlying the major distribution difference among populations is not well understood. However, despite its relative simplicity, the ABO blood group system divides the world's population including both patients and donors into four groups irrespective of origin or creed.

2.10.1 Serology of the ABO Blood Group System

Blood group antigens are molecules expressed on the surface of red blood cells. They consist of proteins and carbohydrates and are often attached to cell surface lipids or proteins. In this regard, the ABO blood group system is comprised of the A, B and H antigens. These antigens define one's individual blood group (A, B, AB or O (H)). In

every individual based on the specific blood group phenotype there are corresponding antibodies, either anti-A or anti-B in serum. Thus type A individuals, express blood group A antigens on the RBC cell surface, and anti-B antibodies; type B individuals, express blood group B antigens on the RBC cell surface, and anti-A antibodies; while type AB individuals, express both blood group A and B antigens on the RBC cell surface, but no antibodies and blood type O type individuals, express blood group neither A nor B, only the backbone H antigen on the RBC cell surface, and both anti A and anti-B antibodies (*Table 2.3*). Anti -A and anti -B antibodies are usually IgM type, and not present in newborns, but appear in the first year of life.

Phenotype	**Antigen (RBC)**	**Antibody (Serum)**
A	A	Anti-B
B	B	Anti-A
AB	A and B	None
O	Neither A nor B	Anti A and Anti-B

Table 2.3: ABO blood group system phenotypes, with corresponding antigen and antibody expression profiles

ABO antigens are one of the oligosaccharide antigens (Storry and Olsson, 2009). The precursor substrate for both A and B antigens is the H antigen. Therefore, the first step in the biosynthesis of the respective A and B antigens is the addition of an L fucose in an alpha (α) 1, 2 linkage to a terminal galactose (Gal) of a common precursor attached to lipids or proteins. This reaction is catalyzed by the enzyme α1, 2 – fucosyltransferase, also known as the H transferase, resulting in the expression of the H antigen. There are

multiple forms of the H antigen known, but the two main structures are Type 1 (Galβ1_3GlcNAcβ-R) and Type 2 (Galβ1_ 4GlcNAcβ -R). The latter is expressed on the surface of red blood cells while the former (i.e. Type 1) is the H antigen expressed in mucosal secretions and in body tissues (Figure 2.7)

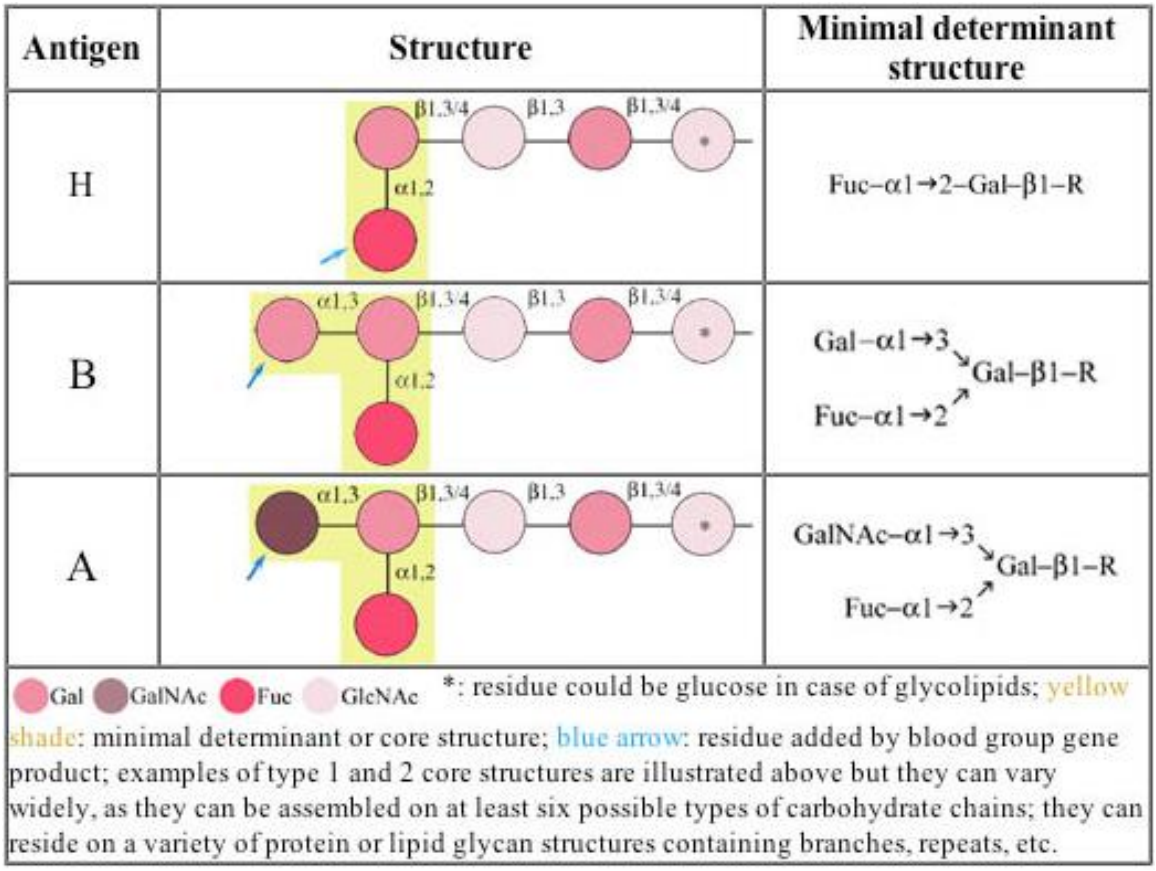


Figure 2.7: Structure of the A, B and H antigen moieties (Source: Blood Group Antigen Mutation Database)

The blood group O phenotype is characterized by expression of blood group H antigen on the erythrocyte surface while modification of the H antigen by the A and/or B glycosyltransferases results in the expression of the A, B and AB antigens. Interestingly the corresponding ABO gene does not encode the respective blood group antigens

directly. This gene, located on the long arm of chromosome 9 (9q34) encodes for the A and B glycosyltransferase that in turn synthesize oligosaccharide epitopes (Storry and Olsson, 2009; Yamamoto *et al*., 1990). The A allele encodes for the A-synthesizing 3-α-N-acetylegalactosaminytransferase and the B allele encodes for the B-synthesizing 3-α-N-galactosaminyltransferase. The resultant gene products i.e. the A and B glycosyltransferases are type II membrane proteins located in the golgi compartment, while soluble forms of the enzymes are found in plasma and other blood fluids and secretions. These enzymes result in the synthesis of the A and B blood group moiteies during normal glycosylation of proteins and lipids in the golgi compartment. Since O allele encodes proteins without glycosyltransferase (O transferase) function, H antigen is the only ABO structure present in blood type O (Storry and Olsson, 2009).

2.10.2 Structure of the ABO Gene Locus

The ABO genes are located in chromosome 9 (9q34.1-q34.2) and consist of 7 exons distributed over an 18 kb region of genomic DNA (Figure 2.8). Exon 7 contains most of the largest coding sequence and Exon 6 contains the deletion found in most O alleles. The exons range in size from 28 to 691 base pairs (Daniels, 2013; Hakomori and Palcic, 2014). The gene locus has three main allele forms, A, B, and O. A and B alleles have seven nucleotide substitutions (297A_G, 526C_G, 657C_T, 703G_ A, 796C_A, 803G_C and 930G_A) of which, four nucleotide substitutions (526C_G, 703G_A, 796C_A and 803G_C) are translated into different amino acid substitutions (Arg526Gly, Gly703Ser, Leu796Met and Gly803Ala). These substitutions determine the specificities of glycosyltransferases (Hakomori and Palcic, 2014).

The A allele encodes A transferase catalyzing the addition of GalNAc residue, and the B allele encodes B transferase catalyzing the addition of Gal residue, respectively, in a α1_3 linkage on terminal Gal of the H antigen. The O allele differs from the A allele by a single nucleotide deletion of guanine (G) at position 261. This deletion causes a frame-shift and results in a loss of transferases activity (Daniels, 2013).

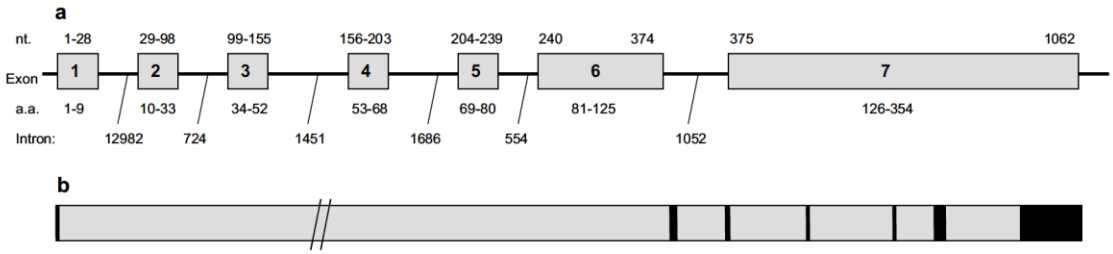


Figure 2.8: Organization of the ABO gene. The ABO gene drawn to scale; exons are black and introns grey (b) (Bahimern et al., 2007)

2.10.3 Physiological Role of the ABO Blood Group System

The ABO blood group antigens are not solely expressed on red blood cells. These antigens are also present in soluble form in body fluids including saliva, breast milk, vaginal and seminal secretions. Their role in transfusion and transplantation medicine has been described over years, based on the serological characteristics of the blood types. In addition to their obvious clinical importance in transfusion and transplantation medicine, there is now extensive knowledge relating to carbohydrate chemistry, enzymology, molecular genetics and structural and evolutionary biology of the ABO blood group system. However, the physiological functions of ABO blood group antigens are still not well understood.

This is because individuals of blood group O individuals express neither A nor B antigens yet, they are perfectly healthy. Therefore this raises a number of questions on the physiological role of the A and B antigens. Furthermore, the variable expression of these antigens as seen in secretors raises a number of questions on what role these antigens play in the context of host disease susceptibilities, especially since mucosal surfaces are often the first point of pathogen host contact. In this regard, a number of associations have been made between ABO phenotypes and increased susceptibility to disease. For example, the blood group O phenotype has been associated with a higher risk of peptic ulcers (Edgren *et al*., 2010). Blood group A with increased *Plasmodium falciparum* malaria disease severity (Wolofsky *et al*., 2012); blood group B, associated with Chagas disease (Teixeira, *et al*., 1987), dengue fever (Kalayanarooj *et al*., 2007) and Human T-lymphotropic Virus-1 (Ayatollahi *et al*., 2008), while type AB individuals have been shown to have an increased incidence of pancreatic cancer (Wolpin *et al*., 2010).

2.11 Minor Blood Group Systems

Although the A, B and H antigens are considered most often, especially clinically with regards to blood transfusion and solid organ transplantation, there are a number of other antigens that may be expressed on the surface of red blood cells. These are considered 'minor' and collectively, comprise the Minor blood group systems. There are a number of studies that have demonstrated a clear correlation between these minor blood group antigens and altered susceptibility to infection. Among these, the Duffy Blood Group Antigen (also known as the Duffy Antigen Receptor for Chemokines, DARC) has been extensively researched.

2.11.1 Duffy Antigen Receptor for Chemokines (DARC)

The Duffy antigen is a transmembrane glycoprotein that spans the red cell membrane seven times. It consists of an extracellular N-terminal domain and a cytoplasmic C-terminal domain. DARC is expressed on venular endothelial cells, cerebellar neurons, and on the surface of red blood cells. Since the discovery of the Duffy Antigen in the 1950s, there has been extensive research on this antigen and its biological roles. The *Fy* gene located on region 1q22-q23 on chromosome 1 of the human genome encodes for DARC (Pogo and Chaudhuri, 2000; Daniels *et al.*, 2005; Rios *et al.*, 2005). The two main alleles described are: FYA, which results in the gene product Fy^a, FYB which results in the gene product Fy^b which lead to four phenotypes Fy (a+b-), Fy (a-b+), Fy (a+b+) and Fy (a-b-). These alleles are codominant (Pogo and Chaudhuri, 2000).

The Duffy glycoprotein consists of 336 amino acids. The extracellular domains of this glycoprotein comprise the N-terminal 60 amino acid domain (ECD1) and three loops

(ECD2-4), comprising of amino acid residues 119 – 129, 189 – 206 and 267 – 288 (Pogo and Chaudhuri, 2000; Daniels *et al.*, 2005; Rios *et al.*, 2005).

The main antigenic epitopes that have been described on DARC: Fy^a, Fy^b, Fy^3 and Fy^6 (Pruenster and Rot, 2006). Fy^a and Fy^b are the main epitopes, relating to a Glycine/Aspartic acid polymorphism at the amino acid residue 42. Fy^6 is a determinant present in Fy^a and Fy^b isoforms. Fy^3 is also an epitope however quite different from Fy^6 mapped to the third extracellular loop of the Duffy glycoprotein (Pruenster and Rot, 2006).

There are two underlying genetic mutations that result in the Duffy negative phenotype: Fy (a-b-). One results from a mutation in the promoter region of the FYB allele. This results in the loss of expression of the Duffy antigen on the surface of the red cells. However, in this case, the Duffy antigen is still produced in other cell types. This is a specific erythroid-specific mutation; quite common among the West African population (100%) and African Americans (70%) (Pogo and Chaudhuri, 2000; Daniels *et al.*, 2005; Rios *et al.*, 2005). It is postulated, since the Duffy antigen is variable expressed among the body tissues, these population, do not express significant amounts of the corresponding antibody Anti-Fy^b. The second less common underlying genetic background is as a result of a point mutation that results in the introduction of a premature stop codon in the coding sequence, resulting in a Fy (a-b-) phenotype. In line with this, Rios *et al.*, 2005 postulate that the truncated Duffy protein fails to be transported to the cell surface and owing to this mutation, the Duffy protein would be absent from all tissues.

Interestingly, the Duffy antigen-negative phenotype, Fy (a-b-), is more prevalent in individuals of African ancestry and exceedingly rare among whites. However, studies have now shown that in Africans whose erythrocytes express the Fy^b gene, antigens are also expressed on the cells of their kidneys, heart, muscles, brain and placenta. Worth noting is that studies have now shown that the Anti - Fy^a is more commonly found in patients of African decent, with the Duffy null phenotype and sickle cell anemics (Pogo and Chaudhuri, 2000; Daniels *et al.*, 2005; Rios *et al.*, 2005).

2.11.2 Duffy Antigen Receptor for Chemokines as a Chemokine Sink

The DARC structure is highly similar to G-protein coupled receptors. However, to date, DARC has not been shown to be a member of this family of receptors. In the early 1990s, the Duffy Antigen was described to be an Interleukin 8 (IL-8) receptor (Darbonne *et al.*, 1991) in addition to other chemokines (Neote *et al.*, 1993 and Horuk *et al.*, 1993), and this lead to its name. DARC binds both CC and CXC chemokines. Chemokines are small, secreted proteins that play a key role in directing leukocyte trafficking. Chemokines traditionally bind and signal through specific G-protein couple receptors with seven transmembrane domains. However, chemokines can also bind to 'silent' receptors known as interceptors including the Duffy Antigen Receptor for Chemokines (DARC) just like classical receptors; this binding has a major impact on chemokine homeostasis (Fra *et al.*, 2003, Shen *et al.*, 2006, Jamieson *et al.*, 2005).

Studies have supported the 'chemokine sink' theory including findings by Dawson *et al.*, (2000) and Shen *et al.*, (2006). Binding of chemokines to erythrocyte DARC, in addition to neutralisation, prevents chemokine diffusion from the blood to the body organs including the lungs, kidneys and other tissues, suggesting that DARC as a chemokine reservoir therefore maintains plasma chemokine levels (Pruenster and Rot, 2006). It is important to note that although there are several associations suggested between DARC expression and plasma inflammatory chemokine levels, intricate details such as how long the chemokines bind, and the function of DARC as a sink, remain unclear (Pruenster and Rot, 2006).

2.11.3 Duffy Antigen Receptor for Chemokines and HIV Infection

DARC expressed on the surface of the red blood cells, as mentioned above, is known to act as a chemokine sink. In association with HIV, recent findings by Weijing *et al.*, (2008) illustrated that erythrocyte DARC mediates binding of HIV-1 to red blood cells; and thereby subsequent transfer of the virions to target cells. However after infection, DARC appears to associate with slower HIV disease progression. DARC expressed on the surface of red blood cells has been shown to influence plasma levels of HIV-associated chemokines such as RANTES (Weijing *et al.,* 2008). RANTES in a pro-inflammatory chemokine with well-known HIV-1 suppressive activities, due to its binding and blocking of the HIV co-receptor CCR5.

Binding to DARC regulates the levels of plasma chemokines associated with HIV pathogenesis. Therefore, it is assumed that expression of DARC may influence the disease progression of HIV/AIDS, where DARC expression slows down disease progression.

It is important to note that although several associations have been suggested between DARC expression and plasma inflammatory chemokine levels, intricate details such as how long the chemokines bind and the function of DARC as a sink remain unclear. The specific mechanism underlying DARC functions remain uncertain, and there is still need to specifically investigate DARC as a variable underlying population-specific differences associated with disease susceptibility. There is need to investigate further the role of DARC in inflammation-associated pathology and malignancy, although this is an area that is highly controversial. My study has been designed to unravel these unanswered questions. We will investigate associations between erythrocyte DARC expression and

susceptibility to HIV infections in a well-established high-risk population of female sex workers in Kenya.

2.12 Secretor Status and *Secretor Fucosyltransferase 2* Gene Polymorphisms

2.12.1 An Overview

The human *Fucosyltransferase 1 (FUT1)* and *Secretor Fucosyltransferase 2* (*FUT2*) genes are responsible for the variable expression of the ABH blood group antigens on red blood cells, and in body tissues, secretions and fluids (Schiff and Sasaki, 1932; Yamamoto *et al.*, 1990a; Yamamoto *et al.*, 1990b). These genes do not directly encode for the ABO blood group antigens. The products of the *FUT1* and *FUT2* genes are glycosyltransferases responsible for the modification of cell surface oligosaccharides resulting in the expression of the A, B and H antigens.

It is estimated, approximately 80% of the global population are 'secretors (*Se*)' and 20% are 'non-secretors (*se*)' (Koda *et al.*, 2001; Ferrer-Admetlla *et al.*, 2009). With regards to the secretor phenotype (both *Se Se* and *Se se*), the *FUT2* gene is expressed in mucin-secreting cells of various mucous membranes and glands, resulting in the secretion of the corresponding blood group antigens into body fluids and epithelial cells (Slomiany and Slomiany, 1978). Based on the underlying genotype profile, the homozygous secretors (*Se/Se*) are known to display more distinct protein expression profiles in comparison to the counterpart heterozygous secretors (*Se/se*). In line with this, the non-secretors are known to express a non-functional form of the protein and/or no protein. Therefore, among non-secretor individuals, ABO antigen expression is restricted to the red blood cells (Koda *et al.*, 2001).

The distribution of the secretor and non-secretor phenotypes, among populations globally has been documented. Global population studies report approximately 80% of the world's population are secretors while 20% are non-secretors (Soejima *et al.*, 2007; Olorunshola

et al., 2013). Furthermore, the secretor phenotype is of particular interest, as there are well-established correlations between secretor status and susceptibility to various bacterial and viral infections. Secretors have been shown to have increased susceptibility to noroviruses. Norwalk virus-like particles (VLPs) were shown to bind to gastro-duodenal epithelial cells from individuals who were secretors but not to cells from non-secretors. This postulation was further confirmed based in *in vitro* studies, where the same VLPs bound to H type-1 thus supporting this association (Huang *et al.*, 2005; Le Pendu *et al.*, 2006). A second investigative study among female study cases with recurrent urinary tract infections, demonstrated secretor status is associated with differences in immunoglobulin concentration. Non-secretors were found to have lower concentrations of serum IgA and IgG; suggesting that this population may have a less effective immune responses in comparison to their non-secretor counterparts thus the underlying increased susceptibility to urinary tract infections (Kinane *et al.*, 1982). The profiling of secretor phenotypes in correlation with oral carriage of Candida albicans, demonstrated that non-secretors were at increased risk of fungal infections (Darwazeh *et al.*, 1990).

In addition, non-secretors have also been found to have increased susceptibility to acute uncomplicated pyelonephritis (Ishitoya *et al.*, 2002), Crohn's disease, one of the major forms of inflammatory bowel syndrome (Dermot *et al.*, 2010), sickle cell disease (Olorunshola and Audu, 2013), Hepatitis B infections (Meo *et al.*, 2010), Type 1 diabetes (Smyth *et al.*, 2011), and the Human Immunodeficiency Virus (HIV) (Ali *et al.*, 2000). In fact, following HIV infection non-secretors have been shown to have slower disease progression (Kindberg *et al.*, 2006). On the contrary, the secretor phenotype has been

associated to increased susceptibility in Helicobacter pylori infections (Linden *et al.*, 2008).

Determination of prevalence rates of secretor and non-secretor phenotypes in a population is therefore important, as this knowledge may inform disease risk and susceptibility patterns; and contribute towards the elucidation of the underlying 'protective' effect conferred by either the secretor and non-secretor phenotypes to infections. In this regard, baseline data on the variable expression of blood group antigens in the context of secretor status in the Kenyan population is not available. This study was therefore designed to determine the frequency of secretors and non-secretors in the Nairobi, Kenya. This study was among a HIV low-risk population of voluntary blood donors and a comparative HIV high-risk population of female sex workers in Nairobi, Kenya.

2.12.2 Secretor Status and the Secretor *Fucosyltransferase 2* (*FUT2*) Gene

Secretor status is defined by the presence of H type 1 blood group antigen in body secretions such as milk and saliva. The H type 1 antigen belongs to the ABO (H) histo-blood-group systems and it is expressed in erythrocyte membranes and in several epithelial tissues (Storry and Olsson, 2009). Individuals of blood groups A, B, or O respectively contain A, B and H antigens made up of complex fucosylated carbohydrates. Fuc α1,2-glycan is the common denominator present in all three ABH antigens because the bone marrow expresses the common H-(fucosyl)transferase (Koda *et al*., 2001; Ferrer-Admetlla *et al*., 2009). The expression of the ABO (H) histo-blood-group systems is regulated by several glycosyltransferases that add monosaccharides to a precursor

molecule in a sequential fashion. In particular, the expression and secretion of ABO antigens in epithelial cells are controlled by secretor type α (l, 2) fucosyltransferase activity, known as the Secretor (Se) transferase. The latter is a product of *Fucosyltransferase 2* gene typically referred to as the *FUT2* gene; a protein-coding gene that constitutes part of a multigene family located on chromosome 19. The α (1,2)- and α (1,3)-fucosyltransferase genes constitute part of a multigene family (Kelly *et al*., 1995). Two human α (1,2) fucosyltransferase genes (*FUT1*and *FUT2*) have been isolated. *FUT1* encodes H type α (1,2) fucosyltransferase (H enzyme) whereas *FUT2* encodes Secretor type α (1,2) fucosyltransferase (Se enzyme) respectively (Koda *et al*., 2001).

In humans the synthesis of soluble A, B and H blood group antigens is determined by the Secretor (Se) (*FUT2*) blood group locus. Glycosytransferases are responsible for the synthesis of oligosaccharides found on the ABO blood group antigens (Kelly *et al*., 1995). They are composed of glycosphingolipids (GSLs) made up of lipophilic portion, ceramide that is attached to the plasma membrane and a variable carbohydrate chain extending out from the cell surface (Suleiman *et al*., 2000). Studies show that among ABH 'secretors' these molecules maybe found in saliva, cell surface of erythrocytes, epithelial cells of the urinary, vaginal, gastro-intestinal, buccal and respiratory tracts and in secretions covering these epithelial cells (Suleiman *et al*., 2000). Based on the expression of the *FUT2* gene, approximately 80% of the global population are 'secretors (*Se*)' and 20% are 'non-secretors (*se*)' (Koda *et al*., 2001; Ferrer-Admetlla *et al*., 2009). With regards to the secretor phenotype (both *Se Se* and *Se se*), the homozygous secretors (*Se Se*) are known to display more distinct protein expression profiles in comparison to the counterpart heterozygous secretors (*Se se*). In line with this, the non-secretors are

known to express a non-functional form of the protein and/or no protein (Koda *et al*., 2001; Ferrer-Admetlla *et al*., 2009).

Although the prevalence of the non-secretor phenotype is similar between populations globally, the point mutations that lead to non-secretors status can differ (Pang *et al*., 2001). There are several single nucleotide polymorphisms of the *FUT2* gene that predict non-secretor status; some may be silent mutations, while others are in coding regions and result in the production of a non-functional form of *FUT2*. Some of the polymorphisms reported to date are population-specific (Koda *et al*., 2001; Ferrer-Admetlla *et al*., 2009). This study will investigate and profile novel *FUT2* polymorphisms within the Kenyan population.

2.12.3 Structure and Function of Secretor Gene

FUT2 is 9.980bp in total length, and it is composed of two exons of 118 and 2.995bp, respectively which are separated by a 6.865bp intron (Admetlla *et al*., 1993) (Figure 2.9). The biosynthesis proceeds from precursors by stepwise addition of monosaccharide units through the action of a set of glycosyltransferases (Marionneau *et al*., 2001). Figure 10 below shows a description of the biosynthesis pathway where type 3 precursor in this case is exclusively found on O-glycans while type 4 is found on glucolipids of the globo and ganglio series respectively. The addition of a fucose in α1,2 linkages gives the H antigens and it is catalyzed by α(1,2) fucosyltransferase (Marionneau *et al*., 2001).

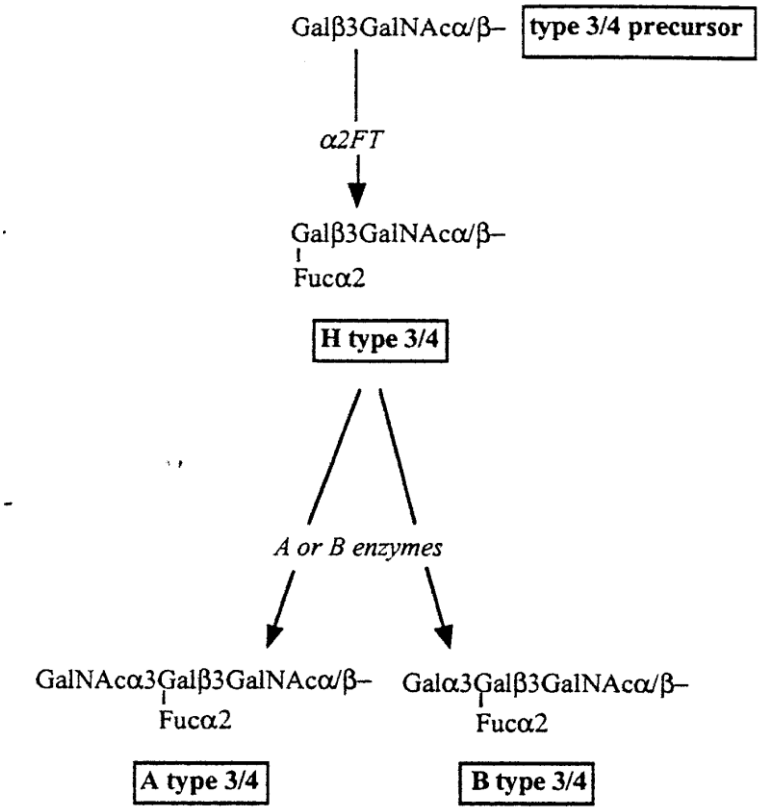


Figure 2.9: Biosynthesis pathway where type 3 precursor in this case is exclusively found on O-glycans while type 4 is found on glucolipids of the globo and ganglio series respectively (Source: Marionneau et al., 2001)

Studies show approximately 20%-30% of humans do not express ABO antigens in saliva, because they are homozygous for the null allele of *FUT2* (Koda *et al*., 2001). This is due to a stop mutation in *FUT2* leading to absence of the enzyme (Kelly *et al*., 1995). These individuals are referred to as "non-secretors" since they are unable to express blood group antigens on mucosal epithelial cells and in secretions (Henry *et al*., 1995) while individuals with "positive secretor status" express the secretor-(fucosyl)transferase which produces the Fuc α1,2-glycan structure (Linden *et al*., 2008).

2.12.4 Secretor *Fucosyltransferase 2 (FUT2)* Gene Polymorphisms

The *fucosyltransferase 2 (FUT2)* gene is composed two exons separated by a single intron and is a highly polymorphic gene. Studies show the first part of the exon constitutes an untranslated coding region, while the second exon codes for a 343-amino acid protein that has extensively been studied. To date, more than 30 single nucleotide polymorphisms (SNPs) have been described across the *FUT2* gene, resulting in various allelic variants of the secretor phenotype (Koda *et al.* 2001; Birney *et al.* 2007; Ferrer-Admetlla *et al.*, 2009). A large number of these polymorphisms are population-specific (Soejima *et al.*, 2007). For instance *Se40, Se375, and Se481* are quite frequent in Xhosa, South-Africa (Liu *et al.*, 1998) and *Se357* and *Se480* in Xhosas, Ghanaians, and Europeans (Kelly *et al.*, 1995; Liu *et al.*, 1998; Soejima *et al.*, 2007).

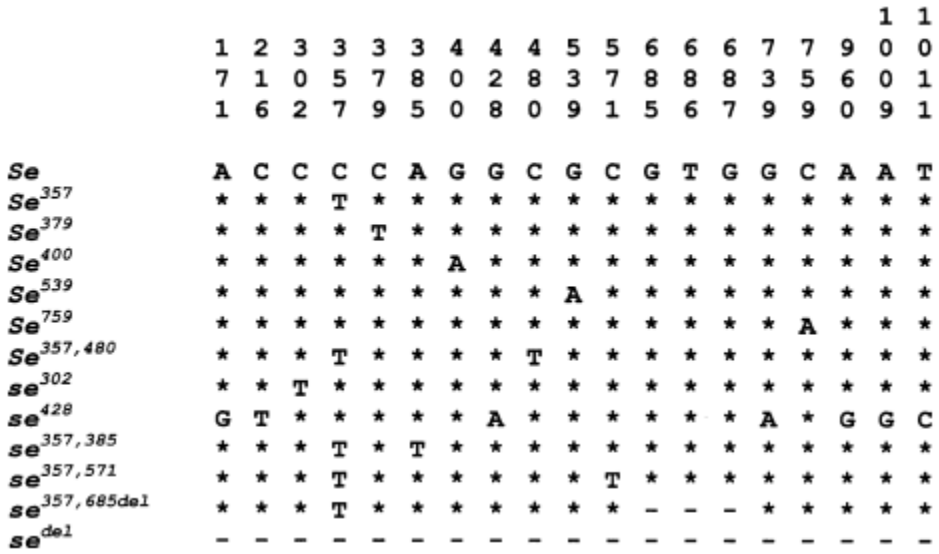


Figure 2.10: Polymorphic base positions in 13 FUT2 alleles.
Functional alleles were indicated as Se and nonfunctional alleles as se. Vertical labels on the top indicate the positions of nucleotides. An asterisk indicates the sequence matched the reference allele (Se) at that position. A - indicates a single base deletion at that position at FUT2 (Pang et al., 2001).

A study done by Pang *et al*., (2001) showed there was little genetic differentiation among the Tamang, Tibetan and Indonesian populations, moderate genetic differentiation between the Uygur and Tamang, and very great genetic differentiation between the Uygur and Tibetan/Indonesian populations. In addition 80% of Caucasians are secretors while the remaining 20% are non-secretors at position 428, this mutation causes a switch of amino acid 143 from a trypyophan to a stop codon resulting in a non-functional protein (Kindberg *et al*., 2006), weak secretor individuals are not yet discovered. In contrast, weak secretor individuals are common among Chinese, Japanese, Polynesians, Australian aborigines and African-Americans (Kindberg *et al*., 2006; Linden *et al*., 2008). The skewed prevalence in secretor phenotypes suggests selection in response to specific types of infections (Linden *et al*., 2008).

2.12.5 Correlation of specific polymorphisms to Infections

The degree of *FUT2* polymorphisms displayed across populations implies that the *FUT2* gene has both negative and positive associations with human health, and therefore has evolved depending on what the heavier selection pressure is for a given population.

Mutations in the second exon of *FUT2* often result in a non-secretor phenotype. Of these mutations, two alleles are considered the most common cause of the non-secretor status: *se428* and *se385* (Ferrer-Admetlla *et al.*, 2009). The non-functional allele *se428* codes for a stop codon at position 143 resulting in a non-functional protein. This allele is responsible for the non-secretor phenotype in several populations including Europeans, Iranians, and Africans (Kelly *et al.*, 1995; Liu *et al.*, 1998; Koda *et al.*, 2001; Ferrer-Admetlla *et al.*, 2009). The latter polymorphism, *se385*, results in a reduction of enzyme activity caused by a missense mutation at position 129 (Ile to Phe). This allele is responsible for the non-secretor phenotype in South East and East Asians (Yu *et al.*, 1995; Henry *et al.*, 1996; Koda *et al.*, 2001; Soejima *et al.*, 2007; Ferrer-Admetlla *et al.*, 2009). Interestingly *se302* is a common non-secretor allele among the Thai and Bangladeshi populations while *se571* is distinct to Samoans (Birney *et al.*, 2007; Soejima *et al.*, 2007). Moreover, one deletion (*se778*), two complete deletions of the coding region (*sedel, sedel2*), and one fusion gene (*sefus*) have been reported (Soejima *et al.*, 2007).

2.12.6 Secretor Status and Disease

There are now well-established correlations between secretor status, *FUT2* gene polymorphisms, and susceptibility to various bacterial and viral infections. Susceptibility to infection by norovirus has also been correlated to the expression profiles of the ABH antigens in the gastrointestinal tract. Norovirus is the most common cause of acute gastroenteritis in humans especially in developing countries (Morrow *et al.*, 2010; Anstee, 2010). Secretor status as a biomarker of norovirus infection susceptibility has been elucidated in a number of studies. In this regard, the null allele, *se428*, has been shown to confer selective protection against the GGII strain of norovirus (Thorven *et al.*, 2005; Larsson *et al.*, 2006). Marionneau *et al.*, (2005) provide supporting evidence to this finding, reporting that heterozygous (*Se/se428*) individuals were more prone to Norwalk-like viral infection than secretor homozygotes (*Se/Se*), whereas non-secretor individuals (*se428/se428*) were relatively resistant to the infection. Thorven *et al.*, (2005) compared susceptibility to gastroenteritis in patients and medical staff involved in hospital outbreaks in Sweden. Their findings demonstrated that homozygous non-secretors were protected from infection. Evidence for greater susceptibility to influenza viruses, rhinoviruses, respiratory syncytial virus, and echoviruses in those with the Secretor phenotype has also been presented (Raza *et al.*, 1991).

Interestingly, a recent study by Morrow *et al.*, (2011) reported findings that during the first weeks of life in neonates, there is a significant increase in *FUT2* gene expression, which directly correlates with an increase in normal gut microbial colonization during infant development. The study postulates that secretor phenotype may provide strong predictive biomarkers of adverse infection outcomes in premature infants.

Few studies have investigated secretor phenotypes and HIV. In Senegalese commercial sex workers, non-secretors were found to have a reduced HIV-1 seroprevalence (Ali *et al.*, 2000). Kindberg *et al.*, (2006) reported a slower HIV-1 disease progression in non-secretors. Confirmation of these findings in a larger population, while determining any potential mechanism involved, is therefore an important research priority.

2.12.7 Postulated Underlying Mechanisms: Secretor Status and Disease

The underlying mechanisms of altered susceptibility to infection associated with blood antigens and their expression profiles (secretor status) are not fully understood, however, there are several hypotheses. Various pathogens including bacteria and viruses express blood group-identical or cross-reactive molecules on their surfaces (Anstee, 2010; Storry and Olson, 2009). These pathogens, have been postulated as probable targets for the corresponding blood group antibodies.

Pathogens including viruses and bacteria have also been shown to selectively bind to the A, B and H blood group carbohydrate molecules via a family of receptor called lectins, which bind specifically to carbohydrate moeities. It is postulated that these pathogens may initiate infection via non-covalent binding to specific mucosal cell surface carbohydrate-binding proteins (lectins). In this way, expression of the blood antigens on mucosal surfaces may place the host at an increased risk of infection (by aiding in pathogen attachment). On the contrary, modification of cell surface carbohydrates at mucosal surfaces by *FUT2* activity may in other instances offer a protective role.

Early studies on viral glycosylation (Arendrup *et al.*, 1991) supported this line of thought. HIV (virion isolates) cultured *in vitro* with peripheral blood mononuclear cells from donors of different ABO groups demonstrated neutralisation by the corresponding ABO antibodies of specific cell isolates. Preece *et al.*, (2002) demonstrated that measles virus, when co-cultured in a system expressing the ABH glycosyltransferases *in vitro*, expressed the corresponding A or B epitopes, according to the enzymes expressed. More recently, Neil *et al.*, (2005) highlight that HIV-1 virions can incorporate ABO blood antigens, both in an artificial trans-infection system, and when primary strains are

propagated in human PBMCs. Further investigations demonstrated that the presence of these antigens sensitizes the virus to the serum of ABO-matched individuals by the action of heat labile complement in conjunction with anti-AB antibodies.

In line with the above, my study has been designed to investigate associations between blood antigens expression profiles, secretor status, *FUT2* gene polymorphisms, and the risk of infection by HIV in a well-established high-risk population of female sex workers in Kenya.

3 MATERIALS AND METHODS

3.1 Study Design

This was a cross-sectional conducted study in Nairobi, Kenya. The target populations were voluntary non-remunerated blood donors recruited at the Nairobi regional blood transfusion centre and female sex workers enrolled in the Pumwani Majengo cohort, recruited from sex worker outreach clinics.

3.2 Study Area

3.2.1 Regional Blood Transfusion Centres

This study was conducted at the Nairobi Regional blood transfusion centre. The target population was voluntary blood donors. The Kenya National Blood Transfusion Services (NBTS) was established in 2001 to collect, test, process and distribute blood and blood components to health facilities and support appropriate blood use in these facilities. The blood transfusion services in Kenya are under a centrally coordinated management structure with a network of regional blood transfusion services. Satellite hospital blood banks have also been set up. These operate under auspices of the NBTS where hospitals do not have easy access to regional blood transfusion services.

3.2.2 Sex Workers Cohorts

Furthermore, this study was conducted in the context of the existing/ongoing University of Manitoba/University of Nairobi collaborative research group activities. The target population were female sex workers from the well established, the Pumwani Majengo sex worker cohort. All self-identified female sex workers within the clinic catchment slum

areas of Pumwani, Kangemi, Kawangware and within the Nairobi Central Business District (CBD), and already enrolled within the Pumwani Majengo cohort and presenting to the target Kawangware, CBD and Majengo clinics were eligible to enroll into the study.

The Pumwani cohort was started in 1983 and has recruited over 2500 participants, with approximately 800 FSWs in active follow up (approximately 22% HIV-infected). The volunteers are primarily recruited from the slum area of Pumwani, Nairobi, Kenya. This cohort is unique in Africa for the duration and depth of follow-up of such mobile, high-risk women. For this unique cohort, sexually transmitted infection (STI) treatment and basic outpatient medical care is available on site. Volunteers complete a detailed questionnaire at enrolment, with a focus on sexual risk behavior (condom use, client numbers, prior genital infections, douching practices, etc.). In addition, they consent to HIV counseling and testing, STI screening and treatment, and blood/female genital tract sampling for immunological studies.

3.3 Study Population

This study enlisted 142 adults of both gender from the Nairobi regional blood transfusion centre. The donors were aged 18 to 65 years. In addition, this study enlisted 280 female sex workers; 90 study cases from the Majengo clinic, 50 cases from Kawangware clinic and 140 from the Nairobi Central Business District (CBD) clinic. The study cases were aged 18 to 65 years of age.

3.4 Sample Size

Sample size calculation was estimated using a standard formula (Hayes and Bennett, 1999) to compare the proportion of individuals with the outcome of interest (*Secretor Fucosyltransferase 2* - *FUT2* gene expression). This was using an estimated power of 80%, a significance level of 0.05, the two populations: secretors versus non-secretors. The calculation was based on the evaluation of the impact of secretor status on HIV susceptibility: 26.7% for the secretors and 17% for the non-secretors as previously described (i.e. non-secretors are less susceptible to HIV (Ali *et al.*, 2000)).

3.4.1 Sample size calculation

The sample size was calculated using a standard formula (Hayes and Bennett, 1999) below:

$$n = (z_{1-\alpha/2} + z_{1-\beta})^2 \frac{\pi_0 (1- \pi_0) + \pi_1 (1- \pi_1)}{(\pi_0 - \pi_1)^2}$$

Where the $f(\alpha, \beta) = (z_{1-\alpha/2} + z_{1-\beta})^2$

Estimated power: 80%

Level of significance: 0.05

Therefore with a power of 80% and level of significance 0.05 $f(\alpha, \beta) = 7.85$

Treatment effect: π_0=17% *versus* π_1=26.7%

Where π_0 and π_1 are the true (population) proportions in the presence and absence of the intervention (secretor status), respectively

Therefore:

$$n = 7.8 \times \frac{17(100-17) + 26.7(100-26.7)}{(26.7-17)^2} = 140$$

This study therefore enlisted 142 blood donors, and 280 female sex workers as summarized in the Figure 3.1 below:

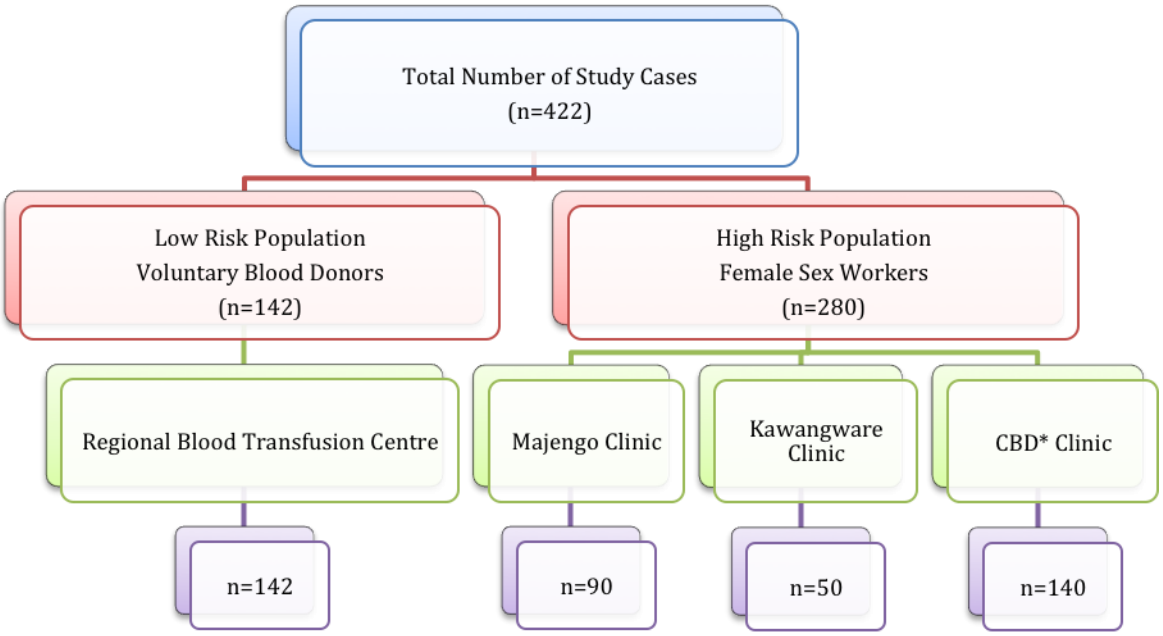


Figure 3.1: Flow Chart of Study Subject Recruitment Per Site

* CBD – Central Business District

3.5 Selection Criteria

3.5.1 Inclusion Criteria for Study Participants

All study participants willing to take part in the study were required to sign the informed consent forms (*Appendix 1: Study Participant Information and Consent Forms*) and also to meet the following conditions:

a) ***For Voluntary Blood Donors:***

- Adults (both male and female) aged between 18 to 50 years of age
- Voluntary, non-remunerated blood donors visiting the Nairobi regional blood transfusion centre (as recruited by the Kenya National Blood Transfusion Services)

Exclusion criteria:

All study participants:

- Who declined to give consent
- Were under the age 18 and over 65 years of age

b) ***For Female Sex Workers:***

Females enrolled in the Pumwani Majengo sex worker cohort

Females aged between 18 and 65 years of age

Females who were not pregnant, based on urine testing

Females not in menses

Females with a uterus and cervix present

Exclusion criteria

All study participants:

- Who did not give consent
- Under the age 18 and over 50 years of age
- Were pregnant, based on urine testing
- With active pelvic inflammatory disease, cervicitis or vaginitis on clinical examination
- Actively menstruating
- Had a prior hysterectomy or ablation of the cervix

3.6 Recruitment of Study Subjects

3.6.1 Regional Blood Transfusion Centres

A project nurse/counselor was appointed to recruit consecutive blood donors visiting the transfusion centres. This was based on a non-coercive approach. All adults, both male and female, aged 18 to 65 years of age, who were visiting the transfusion centres for blood donation were eligible to participate in this study.

As per the National Blood Transfusion Services (NBTS) Policy Guidelines only voluntary blood donors are recruited; these donors are recruited through public campaigns across towns, work places, schools and through the media. NBTS relies heavily on young donors, particularly college and university students. These groups have lower prevalence of HIV and other Transfusion Transmissible Infections (TTIs e.g. Hepatitis B, Hepatitis C and Syphilis) compared to the general population, thereby contributing to the safety and sufficiency of blood and reducing discards from TTI.

Once identified, those who are willing to donate blood join the donation process, which begins with filling of a risk assessment questionnaire. This questionnaire asks about the past medical history of the potential donors and their sexual behaviour. This helps minimize the chances of receiving infected blood and hence reduce wastage of donated blood and equipment. In Kenya, blood is screened after donation; if a sample is found infected and positive for TTIs the sample is discarded. NBTS thus targets donors with the least likelihood of having a TTI to improve safety and minimize discards.

3.6.2 Female Sex Worker Cohorts

This study was conducted at three sex worker outreach program (SWOP) clinics: SWOP

Majengo, SWOP Kwangware and SWOP Central Business District. A project nurse/counsellor at each sex worker clinic recruited consecutive female sex workers visiting the project research clinics. During the clinic visits, there are routine health discussion forums coordinated by the clinic staff at each clinic. It is during these forums that the women, (approximately five (5) women per health discussion forum) are informed and updated regarding any new and/or on-going studies. To address the objectives of this study, a project nurse/counsellor introduced this study; performed the research specific counselling to ensure that all the women understood the study significance, design, objectives and any risks that may have been involved. For the female sex workers interested in participating, the clinic staff then recruited female sex workers consecutively. For the women who consent to participate, once they had fully understood and signed the consent forms; they then proceeded to undergo a brief physical examination by the clinic staff. Based on the outcome and correlation to the outlined selection inclusion/exclusion criteria, the study subjects were recruited. Personnel ensured that recruitment was based on a non-coercive approach. This approach was observed until the target number of study cases was achieved; 90 study cases at the SWOP Majengo clinic, 50 study cases from the SWOP Kawangware clinic and 140 study cases from the SWOP CBD clinic (n=280).

3.6.3 Informed consent

A project nurse/counsellor at each site (regional site and sex worker recruitment clinic) performed research specific counselling to ensure all study participants understood the study significance, design, objectives and any risks that may be involved. The project

nurse hired and appointed at each site obtained informed consent from all study participants, administer survey questionnaire **(*Appendix 1:* *Study Participant Information and Consent Forms*)**.

3.6.4 Screening of Study Cases

For the blood donor based investigations, there was no prior clinical screening involved at the blood transfusion centres. However, all female sex worker study subjects were screened as indicated for sexually transmitted diseases (STDs) including; Gonorrhoea, *Bacterial vaginosis*, and *Trichomonas vaginalis*. Data regarding the concurrent use of contraceptives and sexual behaviour, was taken under consideration at time of cervical sample collection. (The latter information is collected routinely by well-trained project nurses on-site at the Pumwani Majengo site as standard pre-requisite protocol prior to enrolment of any study being undertaken).

3.7 Sample and Specimen Collection

3.7.1 Sampling

Based on the study design and methods, saliva, blood and cervical samples were obtained to perform these studies using safe and well-tested techniques. All specimens were collected from each study participant once informed written consent was obtained. A trained project nurse/counsellor was appointed to collect peripheral blood and saliva samples at each site, and a trained gynaecologist, the female genital tract specimens. Sample analysis was then performed at the KAVI Institute of Clinical Research (KAVI-ICR) laboratories, at the University of Nairobi.

Blood Samples

Voluntary Blood Donors: For the collection of blood samples, two 4ml samples of blood were obtained in EDTA vacutainers from the main sample of blood donated. The participant was not required to undergo a second needle prick following blood donation.

Female Sex Workers: Two 4ml samples of venous blood were obtained in an EDTA vacutainer.

Saliva Samples

For the saliva samples, the study participants were provided with a bottle of clean drinking water to rinse their mouth to remove any debris. A sterile swab, the salimetrics oral swab was placed in the cheek, at the back of the mouth, adjacent to the second molar or at the bottom of the tongue as shown in Figure 3.2. The swab was kept in place for 5 minutes to allow it absorb sufficient saliva. The swab was then removed and placed into a

salimetrics oral swab kit tube. Study participants were then provided some more drinking water to rinse the mouth. There was no risk involved.

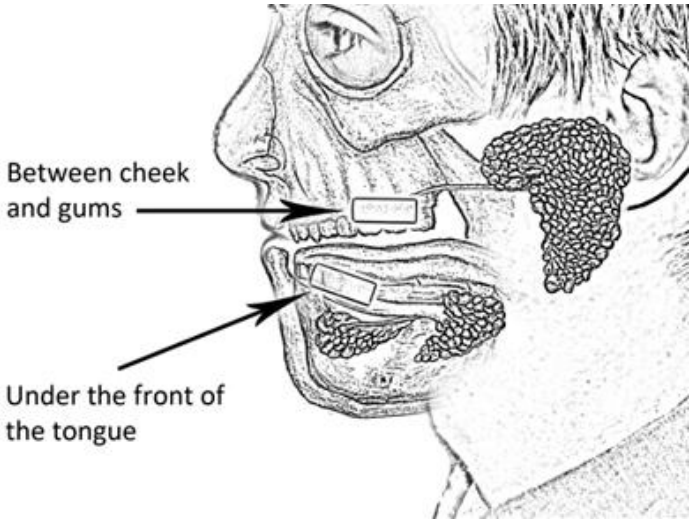


Figure 3.2: Saliva Sampling using Salimetrics KIT (Salimetrics UK)

Female Genital Tract Samples: Female Sex Workers

Data regarding concurrent use of contraceptives and sexual behaviour, was taken under consideration at time of female genital tract sample collection. In case of clinical suspicion or a participant self-report, urine pregnancy testing was performed on-site by the project clinic staff. This was to ensure the study excluded pregnant women.

For the cervical samples, a trained and experienced gynaecologist performed an external genital exam to check the vulva and the opening of the vagina for signs of redness, irritation, discharge, cysts, genital warts or other clinical abnormalities. An internal exam including sampling was then be performed. For the internal exam, the gynaecologist gently inserted a speculum into the vagina. The vaginal walls and cervix were checked for ulceration, discharge, discoloration or other clinical abnormalities.

Samples for further analysis were taken, including:

1. A cervical swab – a sample of cervical mucus was collected using a cotton swab rotated 360^0 in the cervical os and a second swab used to collect secretions from the posterior vaginal fornix. Both swabs were transferred into a single vial containing 5mL of phosphate buffered saline (PBS) which was transported to the laboratory to be tested and cultured for sexually transmitted infections such as *Bacterial Vaginosis* antigen and for screening for *Trichomonas vaginalis* and *Neisseria gonorrhoeae* organisms as indicated by the gynaecologist.

2. Two additional swabs - one vaginal and a second cervical swab as in (1) above were collected, transferred into two separate vials containing 5mL PBS each. These samples were used to screen for blood group antigen secretion.

The samples were then transported to the respective analyses sites, University of Nairobi Institute of Tropical and Infectious Disease (UNITID) and KAVI Institute of Clinical Research (KAVI-ICR) laboratories.

3.8 EXPERIMENTAL APPROACH

The laboratory analyses and investigations were performed at the KAVI-Institute of Clinical Research laboratories, College of Health Sciences, University of Nairobi, Kenya.

3.8.1 Objective 1: To determine ABO, Duffy and Rhesus (D) blood group antigen frequencies among voluntary non-remunerated blood donors in Nairobi, Kenya.

ABO and Rhesus (D) Blood Grouping: ABO and Rhesus (D) blood grouping was determined using commercial antisera kits: murine monoclonal anti-A, anti-B and anti-D antisera (Plasmatec Laboratory Products Ltd., Bridport, Dorset, UK) by the haemagglutination technique on a 2-3% red cell suspension of the blood sample. 100µl of anti-A, Anti-B and anti-D was added into clean test tubes labelled A, B and D, containing 100µl of the 2-3% red cell suspension and incubated at room temperature for 10 minutes. The reaction mix was then centrifuged at 1000rpm for 5 minutes. Red cell buttons were gently resuspended and observed for agglutination macroscopically. Agglutination was interpreted as a positive result and absence of agglutination a negative results. For the ABO blood grouping, negative results were confirmed by reverse typing on the sample serum, and for Rhesus (D) typing, negative results were confirmed using the indirect anti-human globulin test procedure.

Fya, Fyb Blood Grouping: Duffy blood grouping was determined using human monoclonal anti-Fya and anti-Fyb blood grouping reagents (Lorne Laboratories Ltd., Reading, Berkshire, UK) by the Indirect Antiglobulin Technique on a 2-5% red cell

suspension of the blood sample. An equal volume (100µl each) of a 2-5% suspension of washed test red cells prepared in saline was mixed with anti-Fya and anti-Fyb reagents, in two separate tubes, and incubated at 37^0C for 15-30 minutes. Samples were centrifuged at 1000rpm for 5 minutes, and the test cells washed four (4) times in saline. 200µl of anti-human globulin was added to each 'dry' cell button, mixed and centrifuged at 1000rpm for 5 minutes. The cell button was resuspended and agglutination was interpreted as positive. Validity of all negative reactions was confirmed with IgG sensitized red cells by adding a 100µl of Coombs control cells to all negative tubes.

3.8.2 Objective 2: To determine the frequency profiles of secretor and non-secretor phenotypes in the Kenyan population.

Secretor Phenotyping: Secretor status was determined using *Ulex Europaeus* specific anti-H Lectin (Lorne Laboratories Ltd., Reading, Berkshire, UK) to salivary H antigen by the agglutination inhibition technique. 2 ml whole saliva samples collected using sterile salimetrics oral swab kits were centrifuged at 1500 rpm for 15 minutes. The saliva samples were boiled for 10 minutes to denature salivary enzymes, allowed to cool and centrifuged at 1000 rpm for 5 minutes. Equal volumes of the anti-H lectin and sample supernatants (1 ml each) were incubated for 10 minutes at room temperature. An equal volume of a 2-3% O positive red cell suspension was added to each reaction, and centrifuged at 1000 rpm for 5 minutes. No agglutination was interpreted as the secretor phenotype and agglutination indicated the non-secretor phenotype.

3.8.3 Objective 3: To investigate correlations between secretor status and HIV infection rates among a high-risk cohort of female sex workers in Nairobi, Kenya.

Secretor Phenotyping: Secretor status was determined using *Ulex Europaeus* specific anti-H Lectin (Lorne Laboratories Ltd., Reading, Berkshire, UK) to vaginal and cervical H antigen by the agglutination inhibition technique. Swabs collected from the female genital tract were collected using sterile cotton swabs and placed in PBS. These samples were transferred into two separate clean test tubes and centrifuged at 1500 rpm for 15 minutes. Equal volumes of the anti-H lectin and sample supernatants (1 ml each) were incubated for 10 minutes at room temperature. An equal volume of a 2-3% O positive red cell suspension was added to each reaction, and centrifuged at 1000 rpm for 5 minutes. No agglutination was interpreted as the secretor phenotype and agglutination indicated the non-secretor phenotype.

Microbiological Methods: Serological assays were performed on samples as indicated to test for HIV-1 rapid testing (Vironostika), *Bacterial Vaginosis* by microscopy (Gram Stain), *Neisseria gonorrhea* by plate culture using chocolate agar and *Trichomonas vaginalis* by Wet mount.

3.8.4 Objective 4: To investigate correlations between secretor status and HIV infection based on CD4+ T cell counts as a surrogate marker of disease progression.

Immunophenotyping: Samples were analyzed using flow cytometry, the gold standard technique for CD4+ T cell enumeration. Cells were stained with fluorochrome –labeled monoclonal antibodies specific for CD3, CD4 and CD45 (LSR, Becton, Dickinson and Company, San Diego, CA). Samples were acquired on an LSR and data analysis performed using Flow Jo analytical software version 7.2.4 (Treestar, Ashland, OR).

3.8.5 Objective 5: To investigate the genetic variations and underlying polymorphisms among a population of female sex workers in Nairobi, Kenya; with a focus on the *FUT2* se^{428} short nucleotide polymorphism.

Next Generation Sequencing: The *FUT2* gene is located on chromosome 19q13.3 and the protein-coding region is found within the second exon. For secretor genotyping of the entire coding region spanning 1032 bp in length, genomic DNA was prepared from frozen whole blood samples. One 4 ml EDTA tube of blood was obtained, and the samples were frozen at -80^{0}C until sequencing analysis.

DNA extraction: Total DNA was extracted and purified using the QIAamp 96 DNA QIAcube HT kits (Qiagen) intended for automated DNA extraction on the QIAcube HT instrument. Samples were lysed with Proteinase K, and the lysate was loaded onto a QIAamp 96-well plate. A vacuum was established, and during vacuum, DNA was selectively bound to the QIAamp plate membrane as the contaminants passed through. Three wash steps to ensure any remaining contaminants followed this and enzyme inhibitors were removed. Pure DNA was then eluted under vacuum in a single step in

approximately 200ul of Buffer AE. Overlaying the elution buffer with TopElute Fluid enhanced DNA recovery. DNA extraction is summarized in Figure 3.3 below:

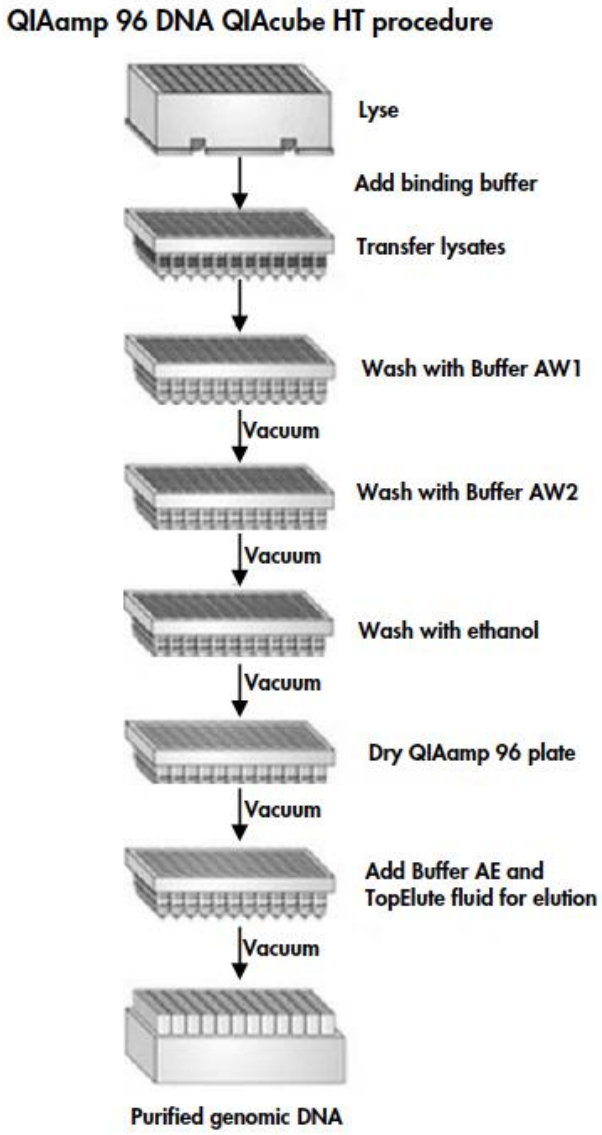


Figure 3.3: Flow chart outlining extraction of genomic DNA

(Source: Qiagen Tehnologies)

Amplification of the coding regions of Secretor FUT2

PCR Primer Design: Appropriate primers were designed for both PCR amplification using the PyroMark Assay Design Software, Version 2.0 as previously described by Ali *et al*., 2000. All primers were then synthesized by Qiagen Sample & Assay Technologies (USA). The PCR Primer sets used are outlined in the Table 3.1 below:

	Pyrosequencing Primers	
Assay 1	TTTCCCATGGCCCACTTC	Forward
	GCAGGGTGATTCTGAAGATGG	Reverse
Assay 2	**CGTTTTCCTCCCCTGACAGC**	Forward
	ATCGTCCACATCCCCCTGAG	Reverse
Assay 3	GTTTTCCTCCCCTGACAGC	Forward
	CCCAGTGCCTTTGATGTTGA	Reverse
Assay 4	**GGGGATGTGGACGATCAA**	Forward
	GCAGGGTGATTCTGAAGATGG	Reverse
Assay 5	CTCAGGGGGATGTGGACGAT	Forward
	AGGCCCCGCAGGAACTTC	Reverse
Assay 6	CTCAGGGGGATGTGGACGA	Forward
	CCCGCAGGAACTTCTGGG	Reverse
Assay 7	GCCCTGGCCAAGATGAAC	Forward
	CCACACTTTTGGCATGACA	Reverse
Assay 8	GGGAGGAGGCCCAGAAGT	Forward
	CAGTCCAGGGCCTGCTGTA	Reverse
Assay 9	**AGGCCCAGAAGTTCCTGCG**	Forward
	GATCCCGAACGTCCCAATG	Reverse
Assay 10	**TGTCCATGTCATGCCAAAAG**	Forward
	AAAATCTTTGGCAGGTGAGC	Reverse

Assay 11	CCTGGTGTCGGGAGAACATT	Forward
	AAGATTTTGAGGAAAGGGGAGTCG	Reverse
Assay 12	**GCCAAAGATTTTGCTCTACTCAC**	Forward
	GAAGGCTGCCTCTGGCTTAAA	Reverse
Assay 13	CACCTGCCAAAGATTTTGCTCTA	Forward
	TCTTGAGGGAGGCAGAGAAGGA	Reverse

Table 3.1: List of primer sets used for amplification of the 1032bp FUT2 region.

PCR Reaction Mix: Sample amplification was based on a quantitative, real time polymerase chain reaction (PCR) assay, using the Rotor-Gene Q (Qiagen). The assay was specifically optimized for Pyrosequencing analysis, to enable highly specific and unbiased amplification of template DNA for sequencing. All reagents were thawed at room temperature and the reaction mixture was set up according to the Table 3.2 below:

Component	**Volume/reaction**	**Concentration**
PyroMark PCR Master Mix (Containing HotStarTaq DNA Polymerase and optimized PyroMark PCR buffer containing $MgCl_2$ and dNTPs)	12.5 µl	
Primer A/B	2.5 µl	0.2 µM
Template DNA	10.0 µl	
Total Volume	**25.0 µl**	

Table 3.2: PCR Reaction mix used during the amplification of the 1032bp FUT2 region

Amplification Cycles: The thermal cycler was programmed according to manufacturer's instructions. The PCR tubes were then placed in the thermal cycler and the cycling program started. The optimized cycling protocol, for 45 cycles, is shown in the Table 3.3 below:

Step	**Duration**	**Temperature**
Initial PCR activation step (Activation of HotStarTaq DNA polymerase)	15 minutes	95^0C
3-Step Cycling Denaturation	30 seconds	94^0C
Annealing	30 seconds	60^0C
Extension	30 seconds	72^0C
Final extension	10 minutes	72^0C

Table 3.3: PCR Amplification Cycle

Confirmation of PCR Products: The amplified PCR products were analysed by electrophoresis in an agarose/Tris EDTA gel containing ethidium bromide and visualised under UV light. After the PCR amplification above, 20 µl of the amplified product was used for subsequent Pyrosequencing analysis.

Sequencing of Amplified PCR Products: PCR products were screened for gene variations using real-time sequence-based Pyrosequencing technology to detect and quantify DNA sequences. The assay was performed using 13 primer sets, so as to achieve genome sequencing of the entire region.

Sequencing Primer Design: Appropriate primers were designed for sequencing reactions using the PyroMark Assay Design Software, Version 2.0 as previously described by Ali *et al*., 2000. All primers were then synthesised by Qiagen (USA). The Sequencing Primer sets used are outlined in the Table 3.4 below:

	Pyrosequencing Primers
Assay 5	AGATGCACAGCACCC
Assay 6	GCAGAACTACCACCTGA
Assay 7	GCTACCCCTGCTCCT

Table 3.4: Sequencing Primer Sets

The sequencing primer set was hybridized to a single stranded, PCR-amplified DNA template. This template was incubated with enzymes and substrates.

The first of four nucleotides, based on an optimized dispensation order of CTGA, was added to the reaction. If the nucleotide was complementary to the base in the template strand, it was incorporated into the DNA strand by the DNA polymerase as shown in Figure 3.4 below:

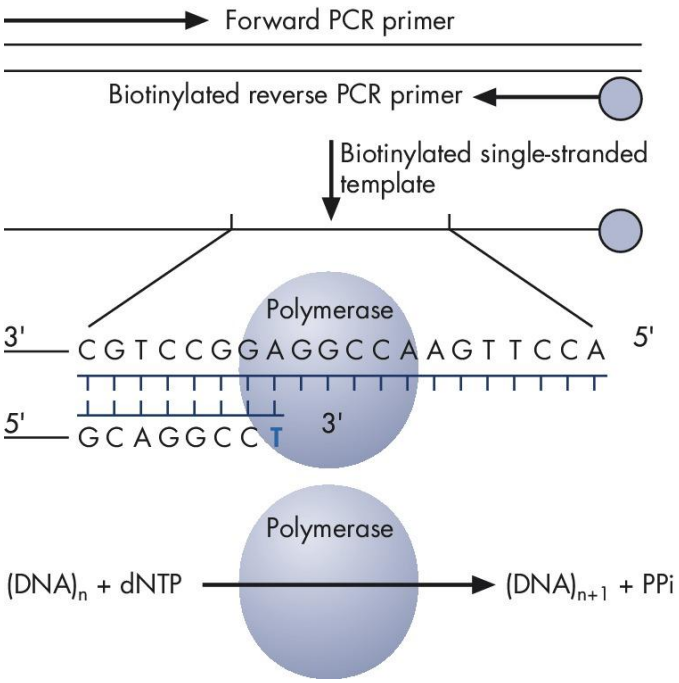


Figure 3.4: Schematic of Polymerase Chain Reaction (Source: Qiagen Technologies)

Each incorporation event was accompanied by release of pyrophosphate (PPi) in an equimolar quantity to the amount of nucleotide incorporated. ATP sulfurylase quantitatively converted PPi to ATP in the presence of adenoside 5' phosphosulfate. This drove the conversion of luciferin to oxyluciferin by luciferase, generating visible light in amounts proportional to the amount of ATP. Light was detected using charged coupled devices (CCDs) and seen as a peak in the Pyrogram®. Each light signal was proportional to the number of nucleotides incorporated as shown in Figure 3.5 below:

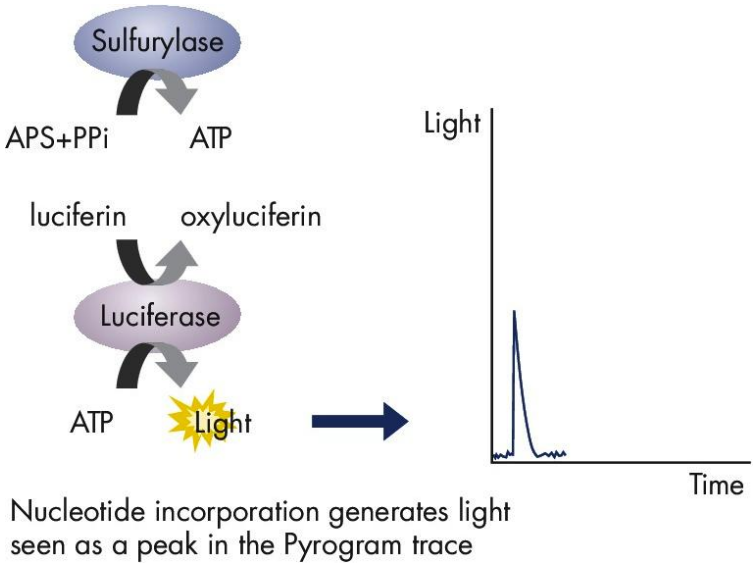


Figure 3.5: Nucleotide incorporation results in generation of the peak as above (Source: Qiagen Technologies)

Apyrase, a nucleotide-degrading enzyme, continuously degraded unincorporated nucleotides and ATP. When degradation was complete, another nucleotide was added. As the process continued, the complementary sequence was built up and the nucleotide sequence was determined from the peak in the Pyrogram as shown in Figure 3.6 below.

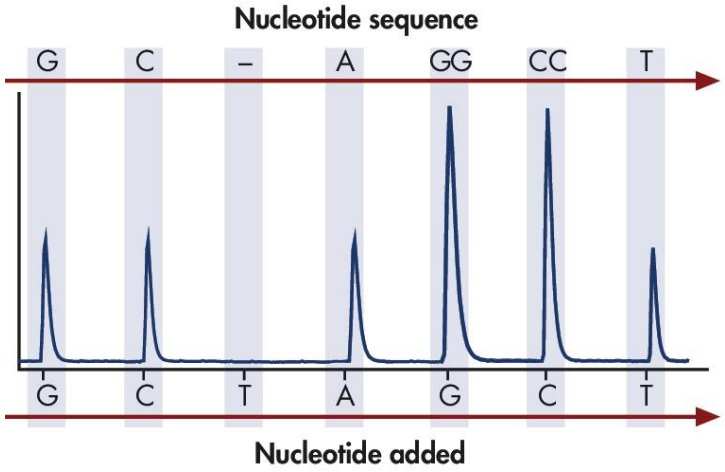


Figure 3.6: Pyrogram generated during Sequencing (Source: Qiagen Technologies)

On completion, the run was saved on an external drive and the respective Pyrograms per sample well were analyzed using the PyroMark Q24 Software.

Analysis of the run

The processed run file was transferred from the external drive (USB stick) to the computer running the PyroMark Q24 Software. The data per well was analyzed in 'SQA' mode. By selecting an analyzed well, the corresponding Pyrogram was displayed and the well information (including analysis warnings) were displayed as demonstrated in Figure 3.7 below:

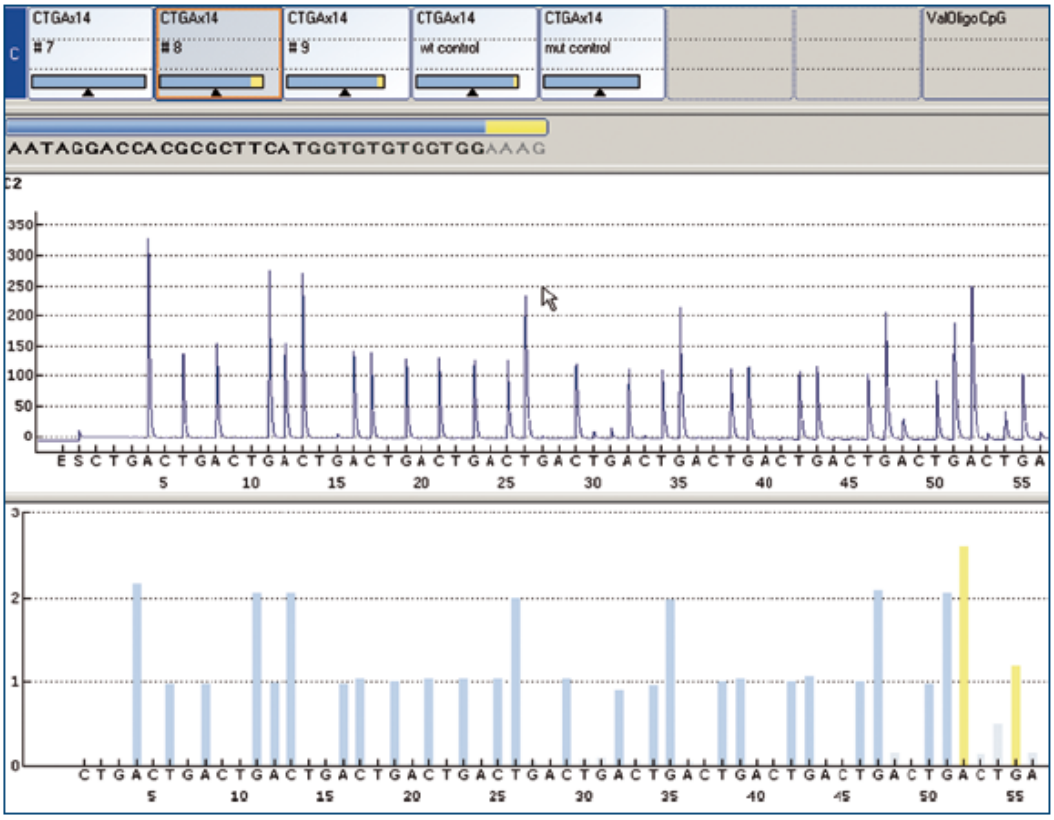


Figure 3.7: Quality Analysis Checks During Sequencing Run; Blue (Passed), Yellow(Check), Red (Failed) (Source: Qiagen technologies)

3.9 Quality Assurance

Quality assurance was performed to ensure that all the tests and analyses from both the UNITID and KAVI laboratories were correct, and relevant to the study investigations. To achieve this, all laboratory work was under skilled management with critical supervision of the work including quality assurance programs and practices in place.

Quality Assurance Programs: The current programs in place at both KAVI-ICR and UNITID laboratories cover all the respective laboratory activities aimed at achieving the required standard of analysis. The various aspects of quality assurance considered include internal quality assurance and control, external quality assurance, standardization and optimization of methods and method proficiency surveillance of all analytic procedures.

Internal Quality Assurance: The internal quality assurance assayed the precision and reproducibility of the assays. This will include pre-analytic and analytic controls.

Pre-Analytic Controls: Standardized specimen collection and specimen transit to the respective sites of analyses. For each sample collected there was a pre-ordered test requested and recorded. There were also quality assurance checks already in place at the laboratories concerned with the processes from specimen collection, labelling, delivery and storage prior to any analytical analyses.

Analytic Controls: There are currently standard operating procedures written for the

standard and routine specimen collection and handling, test performance, reporting of results, quality control and quality assurance. Critical values, reporting critical results and corrected reports are all defined and written. There is also adequate space and facilities available. Local and National regulations and guidelines are adhered to for infection control, hazardous/infectious waste disposal adhered to and documented.

Furthermore:

- All the technical procedures include the test principle, specimen required, equipment/reagents needed, directions for performing the tests, sources of error, interpretation of results (includes criteria for repeating/referring specimens for further review), reporting protocol and references.
- Test kits and reagents are correctly labelled, stored at the proper temperatures and used within the expiration dates
- Documentation that equipment/procedure calibration is done upon implementation of the instrument or method, as required by manufacturer and when controls show trends, shifts or are out of limits
- Worksheets and printouts are retained for the most recent 2 - 5 years
- Documentation of new instrument/test validation studies are available
- Reference ranges established/verified for control materials and documentation available
- All lab reference values are valid for their respective test methods and relevant to the local and study population
- Reference books/atlases are available for identification of unknowns

- Equipment maintenance is performed as appropriate and documented
- Corrective actions are all documented
- Documentation that reagents prepared/stored and used at proper temperatures

Records, Training and Maintenance: Efficiency of recording and reporting of results, maintenance and control of equipment and apparatus, staff training, administrative procedures, management, protection of laboratory staff against health risks and hazards when handling specimens and equipment. Unique patient and sample codes, date (and time, if appropriate) results reported, unit of measure for each value, specimen source and limitations and normal ranges. Equipment function checks kept for up to two years and maintenance records for life of the instrument. Lot numbers, expiration dates of kits, reagents, controls, calibrators, and standards kept for 2-5 years. Requisitions, quality control results, proficiency testing data; biannual verification of accuracy of tests, quality assurance activities.

External Quality Assurance: The external quality assessment will be concerned with the inter-laboratory and inter-method or inter-instrument harmony. This will entail the regular comparability of results wheresoever the test is performed. There are currently annual external quality assurance surveys. These ensure that all assays being performed at the laboratories are standardized and optimised. For each selected methods there will be instrument calibration and equipment checks.

Accreditation

The KAVI-ICR laboratories have already attained accreditation under the stringent Good Clinical Laboratory Practice (GCLP) scheme. Annual systematic study audits are carried out to determine if the laboratory maintains a satisfactory level of good practice with effective application of quality-related issues in accordance with any regulatory requirements.

3.10 Data Management and Statistical Analyses

Laboratory data was entered in a database using Micorsoft Excel software and exported to Scientific Package for Social Science (SPSS) version 20 for statistical analyses. Frequency distribution of the dependent categorical variables were compared by the Chi-square test. Chi-square (χ^2) tests were performed to compare frequency distribution of categorical variables, with odds ratios (ORs) used as measures of association. The level of significance for all the analyses was set at a value of P = 0.05 at 95% Confidence Interval (CI). Multivariate logistic regression analysis was used to analyse data for the association between HIV infection and the secretor phenotype, adjusted for potential confounding variables. The presence of confounding was determined by comparing the ORs obtained from logistic regression models before and after addition of the covariates being evaluated. Stratified logistic regression analyses were performed to assess the relationship between HIV infection and the secretor phenotype with female sex workers; comparing HIV infected and HIV un-infected. In addition, multivariate logistic regression analysis was used to analyze data for the association between HIV infection and the secretor phenotype against CD4+ T cell counts, adjusted for potential

confounding variables. The presence of confounding was determined by comparing the ORs obtained from logistic regression models before and after addition of the covariates being evaluated. Stratified logistic regression analyses were performed to assess the relationship between HIV infection and the secretor phenotype among the female sex workers; comparing HIV infected and HIV un-infected.

4 RESULTS

The study began with a baseline profile of the frequencies of ABO, Rhesus (D) and Duffy (Fy^a, Fy^b) blood group antigen phenotypes; followed by the determination of secretor phenotype frequencies among both the blood donors and female sex workers. Among the female sex workers secretor phenotypes were correlated to HIV infection rates, further based on CD4+ T cell counts as a surrogate marker of disease progression. The last objective involved the screening for the FUT2 se^{428} polymorphism among the same cohort of female sex workers. The results of this study are presented below.

4.1 Socio-Demographic Characteristics of the Study Population

4.1.1 Gender

This study recruited 422 adults in Nairobi, Kenya. The study population comprised 142 voluntary blood donors (106 males and 36 females) and 280 female sex workers.

4.1.2 Nationality

Majority of the study participants were from Kenya, however there were some non-Kenyan nationals who were enlisted at the time of the study as outlined in Table 4.1 below:

Nationality	**Blood Donors**	**Female Sex Workers**
Kenyan	134 (94.4%)	249 (88.9%)
Tanzanian	-	7 (2.5%)
Congolese	-	1 (0.3%)
Ugandan	-	2 (0.7%)
Arab	2 (1.41%)	-
Declined to Disclose	6 (4.22%)	21 (7.5%)
Totals	**142**	**280**

Table 4.1: Nationalities among the study participants.

4.1.3 Age Stratification among Study Participants

The age range for all study volunteers was 18-65 years with a mean age of 33.01 years (SD, 9.3 years), as outlined in Table 4.2. These were 142 voluntary blood donors and 280 female sex workers (n=422). **The percentage was calculated based on the group totals i.e. (No./142100) for blood donors, (No./280100) for female sex workers and (No./422**100) for the totals respectively.

	Blood Donors		**Female Sex Workers**		**Total**	
Age (Years)	**No.**	**%***	**No.**	**%***	**No.**	**%***
18-24	60	42.2	29	10.4	89	21.1
25-29	39	27.5	48	17.1	87	20.6
30-34	17	11.97	55	19.6	72	17.1
35-39	10	7.04	58	20.7	68	16.1
40-44	10	7.04	28	10.0	38	9.0
45-49	3	2.1	34	12.1	37	8.8
50+	3	2.1	28	10.0	31	7.3
Totals	**142**		**280**		**422**	
Means	**27.97 (SD, 7.7 years)**		**36.1 (SD, 9.3 years)**		**33.01 (SD, 9.3 years)**	

Table 4.2: Age distribution of the study participants.

Majority of the blood donors were under the age of 30 (69.7%), while the female sex workers were mostly above the age of 30 years (72.5%) as shown in Figure 4.1 below. This difference was statistically significant (p=0.0001).

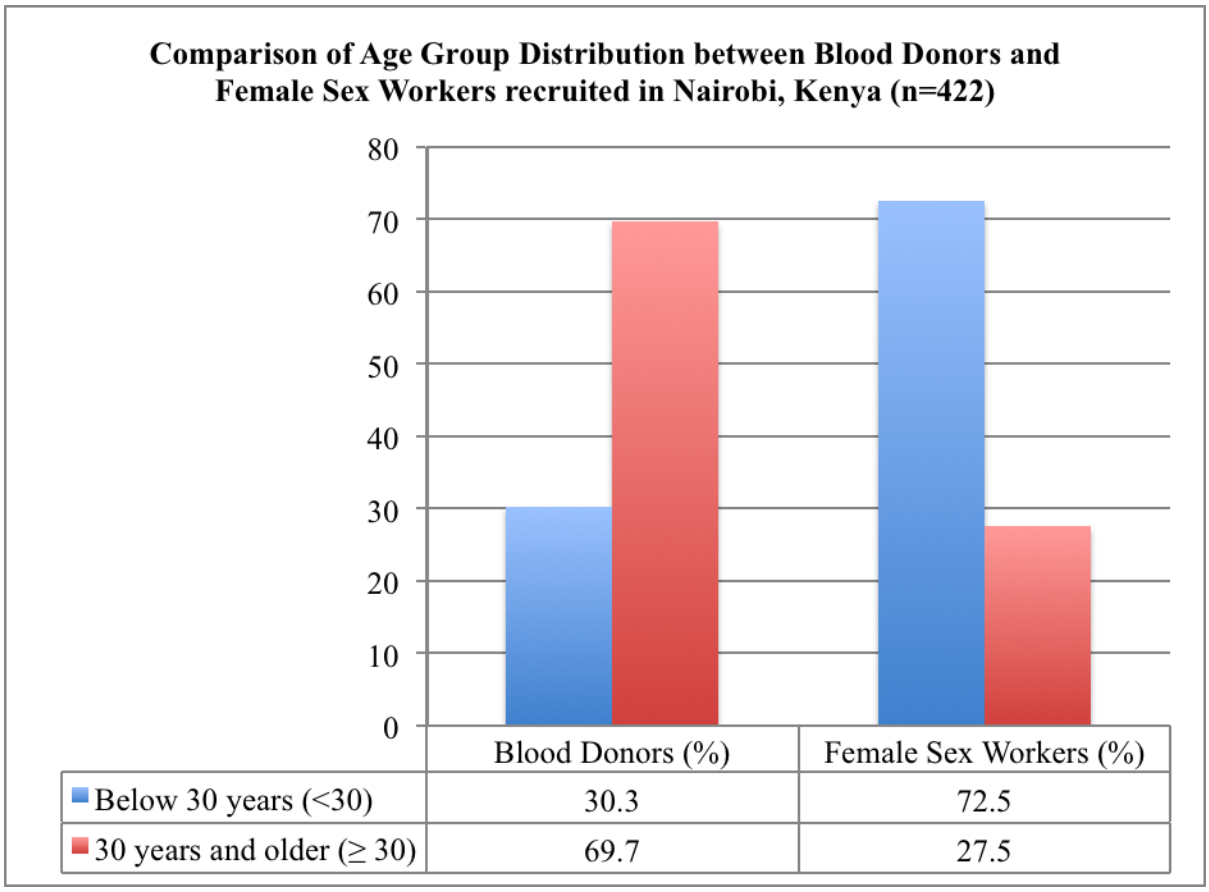


Figure 4.1: Age group distribution of the study participants.

4.1.4 Female Sex Worker Cohort Characteristics

Majority of the female sex workers enlisted had practiced sex work for more than 5 years, with the mean duration at 8.75 years. The main contraceptive used was the male condom, while most of the women reported to have children, mainly 2-5 children per participant. These characteristics are outlined in Table 4.3 below.

POPULATION CHARACTERISTIC	NUMBER OF STUDY CASES (n)
Duration in Prostitution	
0-2 years	59 (21.1%)
3-5 years	90 (32.1%)
6-8 years	39 (13.9%)
9+ years	92 (32.9%)
Mean Duration in Prostitution	8.75 years (SD, 7.9 years)
Contraceptive Use	
Oral Contraceptive Pill	20 (7.1%)
Intrauterine Contraceptive Device	5 (1.8%)
Depo-Provera	39 (13.9%)
Tubal Ligation	4 (1.4%)
Male Condom	196 (70%)
Female Condom	1 (0.4%)
Diaphragm	2 (0.7%)
Other	13 (4.6%)
Number of Children	
0-2	207 (73.9%)
3-5	69 (24.6%)
6+	4 (1.4%)

Table 4.3: Socio-demographic characteristics of the female sex workers.

4.2 ABO, Duffy and Rhesus Blood Group Antigen Phenotypes

ABO blood group phenotype distribution was O>A>B>AB among both the blood donors and female sex workers. Collectively, blood group O were a majority comprising 199 (47.2%), A 120 (28.4%), B 84 (19.9%) and AB 19 (4.5%) respectively with 408 (96.7%) Rhesus (D) positive cases. Duffy positive phenotypes were reported in 2 (0.47%) of the

study participants, both female sex workers. The frequencies of the respective blood group phenotypes are shown in the Table 4.4.

	Blood Donors		**Female Sex Workers**		**Total**	
Blood Group	**No.**	**%**	**No.**	**%**	**No.**	**%**
ABO						
O	58	40.8	141	50.4	199	47.2
A	47	33.1	73	26.1	120	28.4
B	30	21.1	54	19.2	84	19.9
AB	7	4.9	12	4.3	19	4.5
Rhesus (D)						
Positive	135	95.1	273	97.5	408	96.7
Negative	7	4.9	7	2.5	14	3.3
Duffy (Fya, Fyb)						
Fya^{+} Fyb^{-}	0	0	1	0.3	1	0.2
Fya^{-} Fyb^{+}	0	0	1	0.3	1	0.2
Fya^{-} Fyb^{-}	142	100	278	99.3	420	99.5

Table 4.4: ABO, Rhesus (D) and Duffy blood group distribution of the study participants.

Further stratification, based on the ABO and Rhesus (D) phenotypes confirmed > 95% of the blood donors were Rh positive as outlined in Table 4.5 below.

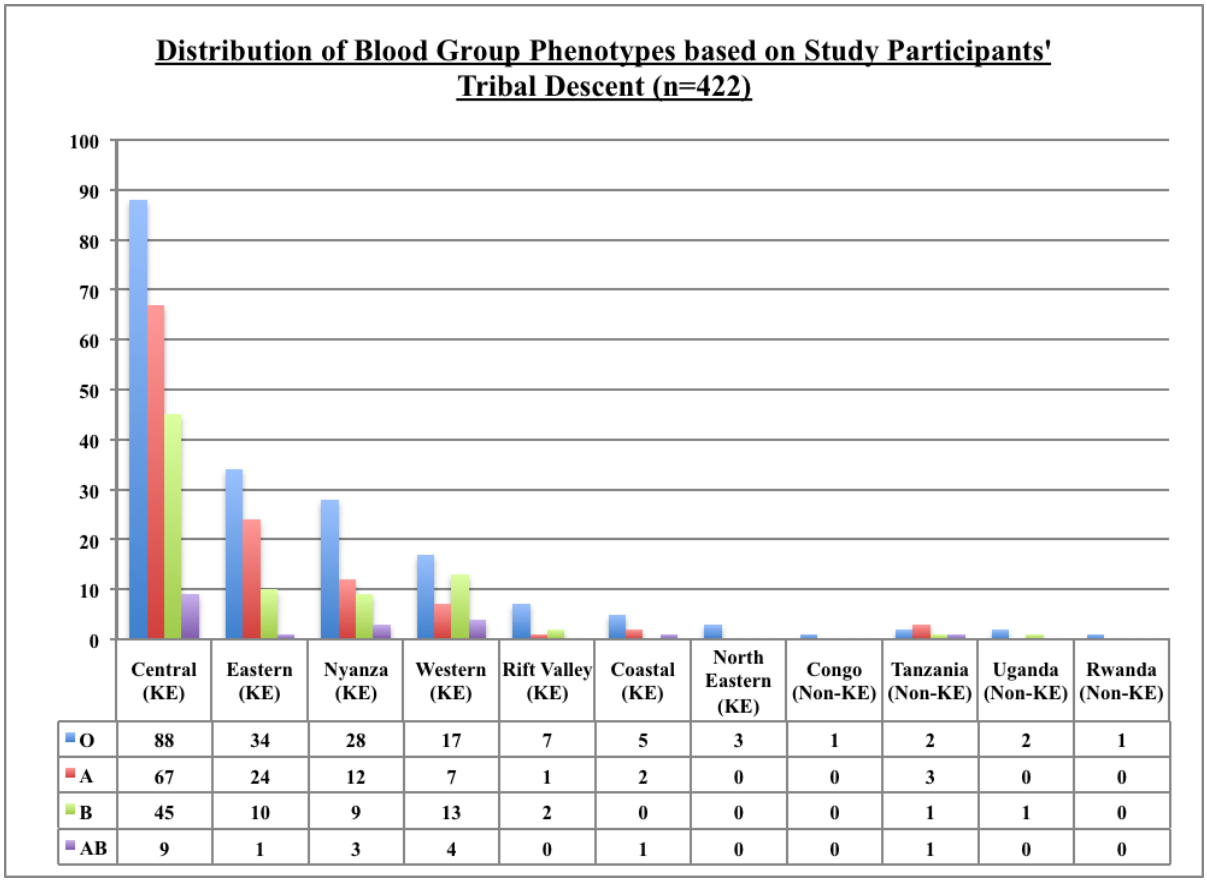
	Blood Donors		Female Sex Workers		Total	
Blood Group	**No.**	**%**	**No.**	**%**	**No.**	**%**
O						
Positive	55	94.8	137	97.2	**192**	**96.5**
Negative	3	5.2	4	2.8	**7**	**3.5**
A						
Positive	45	95.7	71	97.3	**116**	**96.7**
Negative	2	4.3	2	2.7	**4**	**3.3**
B						
Positive	28	93.3	54	100	**82**	**97.6**
Negative	2	6.7	0	0.0	**2**	**2.4**
AB						
Positive	7	100	11	91.7	**18**	**94.7**
Negative	0	0.0	1	8.3	**1**	**5.3**

Table 4.5: Distribution of Rhesus (D) phenotypes between ABO phenotypes of the study participants.

4.3 Regional Distribution of ABO Phenotypes

The study participants were all recruited from Kenya, however, there was further stratification based on individual ancestral descent. 392 (92.9%) were from Kenya and of these, study participants from Central region represented a 209 (49.53%) majority, while study cases from North-Eastern region represented a 3 (0.71%) minority. The non-Kenyan study participants represented a 2.8% minority from East and West Africa (Figure 4.2). In Central, Eastern, Nyanza, and Coastal regions of Kenya the blood group phenotype distribution patterns were O>A>B>AB, while in Western and Rift valley regions, the distribution patterns were O>B>A>AB. 18 (4.27%) of the study participants

declined to disclose their nationality and/or tribe.



	Central (KE)	Eastern (KE)	Nyanza (KE)	Western (KE)	Rift Valley (KE)	Coastal (KE)	North Eastern (KE)	Congo (Non-KE)	Tanzania (Non-KE)	Uganda (Non-KE)	Rwanda (Non-KE)
O	88	34	28	17	7	5	3	1	2	2	1
A	67	24	12	7	1	2	0	0	3	0	0
B	45	10	9	13	2	0	0	0	1	1	0
AB	9	1	3	4	0	1	0	0	1	0	0

Figure 4.2: Regional distribution of blood group phenotypes.

4.4 Secretor and Non-secretor Status

Saliva testing showed that among the blood donors, 121 (85%) were secretors and 21 (15%) were non-secretors, while among the female sex workers, 212 (76%) were secretors and 68 (24%) non-secretors. In total, 333 (78.9%) of the study population were secretors and 89 (21.1%) were non-secretors. (Figure 4.3).

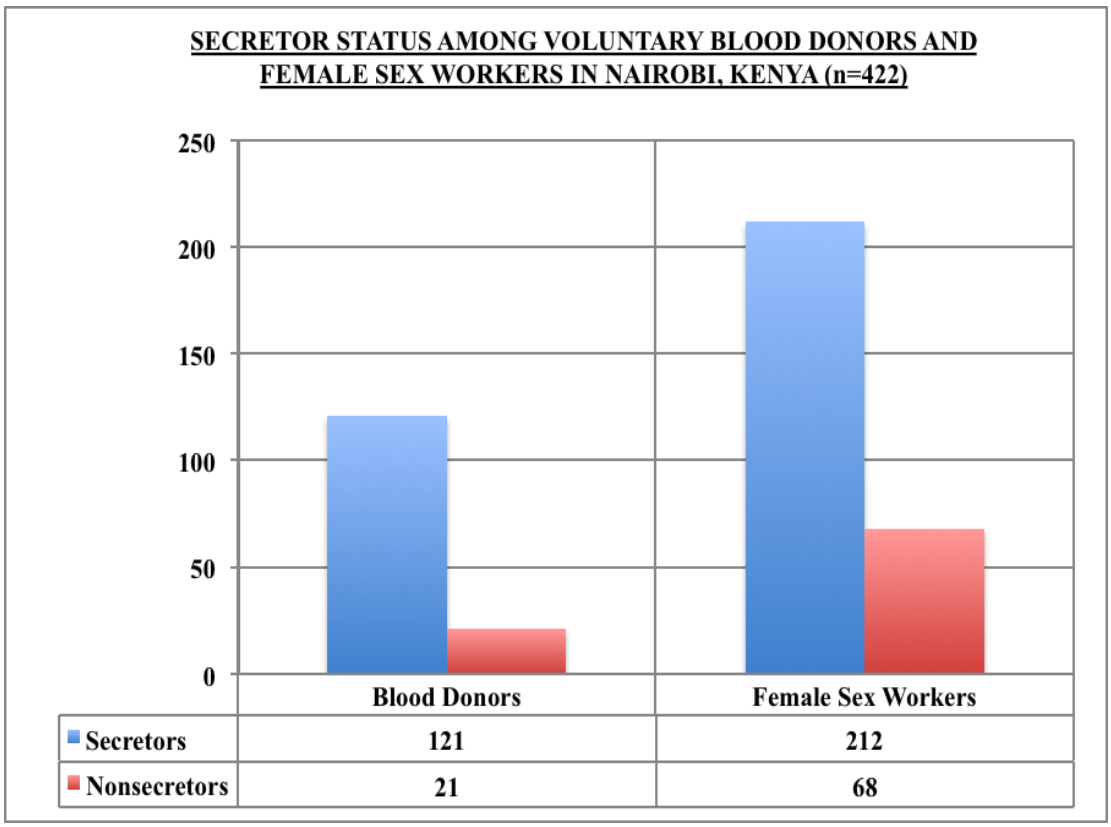


Figure 4.3: Secretor status of voluntary blood donors (n=142) and female sex workers (n=280) in Nairobi, Kenya.

4.4.1 Secretor Status and Gender

The study participants comprised 142 blood donors (106 male and 36 female), and 280 female sex workers. Of the 106 male study subjects 87.7% were secretors and 12.2% were non-secretors, while among the 316 female study participants 77.5% were secretors and 22.5% were non-secretors. Statistical analysis of the proportions indicated the secretor phenotype was more frequent among the male study subjects (87.7%) in comparison to the female study cases (75.9%). There was a significantly higher proportion of secretors among the male study participants (p=0.027) (Table 4.6).

Gender	Male	Female	P Value
Secretors	93 (87.7%)	240 (75.9%)	P=0.027
Non-secretors	13 (12.2%)	76 (24.1%)	
Total	**106**	**316**	

Table 4.6: Distribution of secretor status prevalence based on gender.

4.4.2 Secretor Status and ABO Phenotypes

Based on the ABO blood group phenotypes, blood group O secretors were a majority: n=163 (81.9%) and blood group AB secretors the least frequent: n=14 (73.7%). Collectively, the proportion of secretors versus non-secretors in each blood group ranged from 73% - 82% for the secretors and 18% - 26% for non-secretors as shown in Table 4.7.

	Blood Donors		**Female Sex Workers**		**Total**	
Blood Group	**No.**	**%**	**No.**	**%**	**No.**	**%**
O						
Secretors	51	87.9%	112	79.4%	163	81.9%
Non-secretors	7	12.1%	29	20.6%	36	18.1%
A						
Secretors	40	85.1%	52	71.2%	92	76.7%
Non-secretors	7	14.9%	21	28.8%	28	23.3%
B						
Secretors	25	83.3%	39	72.2%	64	76.2%
Non-secretors	5	16.7%	15	27.8%	20	23.8%
AB						
Secretors	5	71.4%	9	75.0%	14	73.7%
Non-secretors	2	28.6%	3	25.0%	5	26.3%

Table 4.7: Frequency of secretor status in A, B, AB and O blood group individuals.

However, when comparing the individual blood groups, there was no significant

difference in the incidence of secretors per blood group phenotype as shown in Figure 4.4 below:

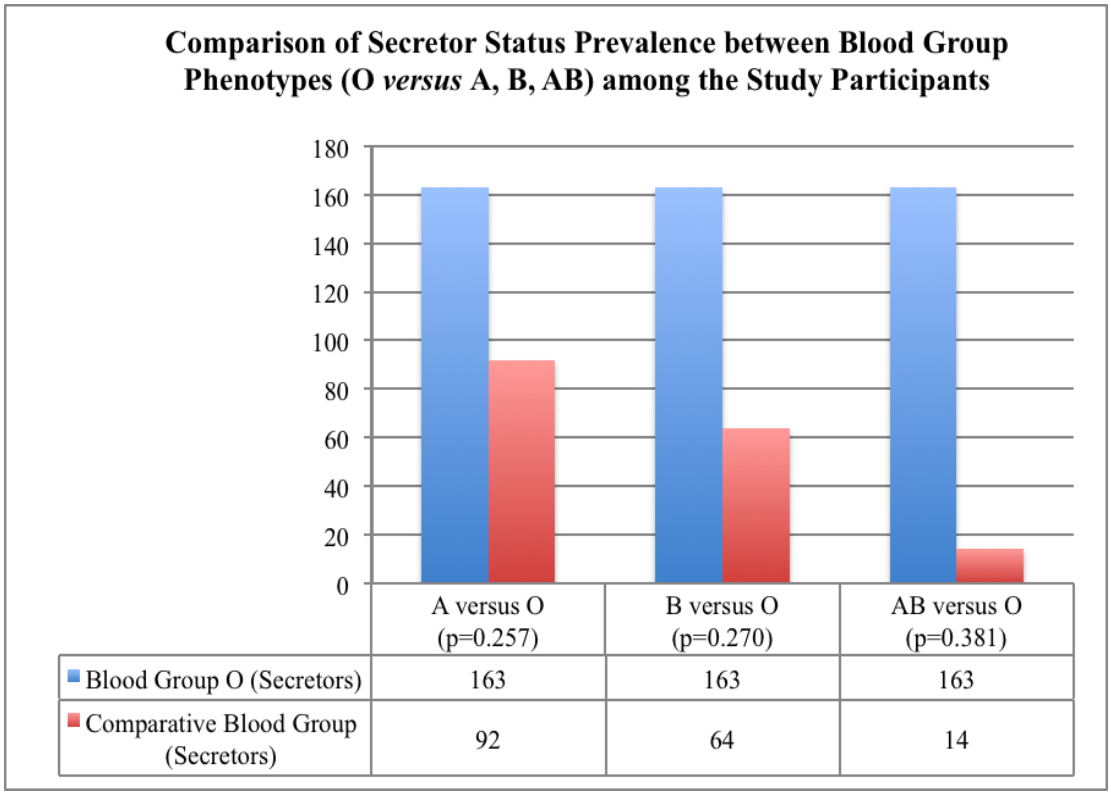


Figure 4.4: Comparison of Secretor status prevalence between blood group phenotypes (O, n=163; A, n=92; B, n=64; AB, n=14)

4.5 HIV Screening:

The female sex workers were screened for HIV-1. 92 (32.9%) were HIV positive and 188 (67.1%) HIV negative (Figure 4.5).

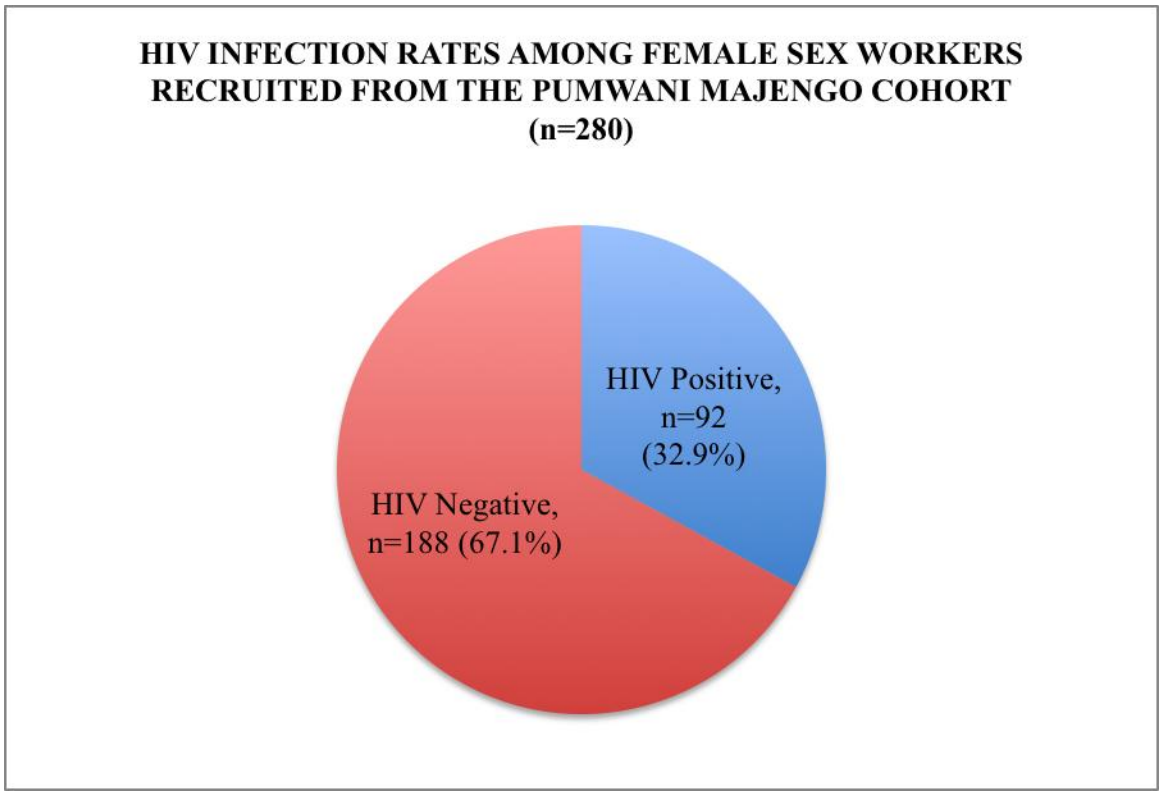


Figure 4.5: HIV Infection Rates of female sex workers recruited from the Pumwani Majengo Sex Worker cohort (n=280)

4.5.1 Medical History:

Study participants were required to complete a health assessment questionnaire during their clinic visits, outlining their medical history over the past one year. Of these 55 of the 92 HIV infected women (59.8%) were on antiretroviral therapy, and 17 of the 280 study cases (6.1%) reported a history of tuberculosis treatment.

4.5.2 Secretor Phenotypes and HIV Infection Rates:

As shown in figure 4.3, above, of the 280 female sex workers, 212 (75.7%) were secretors and 68 (24.3%) non-secretors. This was consistent in swabs obtained from vaginal and cervical sampling sites. On evaluation of HIV status, the proportion of secretors was significantly higher among women with HIV infection (77/92 = 83.7%)

than among the HIV un-infected women (135/188 = 71.8%) p=0.029 (Table 4.8).

Phenotype	HIV Positive	HIV Negative	P Value
Secretors	77 (83.7%)	135 (71.8%)	p = 0.029
Non-secretors	15 (16.3%)	53 (28.2%)	
Totals	92	188	

Table 4.8: Distribution of secretor and non-secretor phenotypes of HIV infected and HIV uninfected female sex workers recruited from the Pumwani Majengo Sex Worker Cohort (n=280)

4.5.3 Blood Group Phenotypes and HIV Infection Rates:

Among the secretors, when stratified based on the individual ABO phenotypes (Table 9), the incidence of HIV infection was higher among blood group A secretor individuals (26/52 = 50%), in comparison to blood group B (12/39 = 33.3% : p = 0.066), AB (3/9 = 33.3% : p=0.355), and this difference was statistically significant when compared to blood group O secretors (36/112 = 32.1% : p = 0.028) (Figure 4.6). There was a significant difference in the incidence of infection between A secretors and O secretors.

Phenotype	A (n=73)	B (n=54)	AB (n=12)	O (n=141)	Totals
Secretors (Totals)	52	39	9	112	212 (75.7%)
	HIV+ = 26 (50%) **HIV- = 26** (50%)	**HIV+ = 12** (30.8%) **HIV- = 27** (69.2%)	**HIV+ = 3** (33.3%) **HIV- = 6** (66.7%)	**HIV+ = 36** (32.1%) **HIV- = 76** (67.9%)	
Non-secretors (Totals)	21	15	3	29	68 (24.3%)
	HIV+ = 4 (19.04%) **HIV- = 17** (80.96%)	**HIV+ = 3** (20%) **HIV- = 12** (80%)	**HI V+ = 2** (66.7%) **HIV- = 1** (33.3%)	**HIV+ = 6** (20.7%) **HIV- = 23** (79.3%)	

Table 4.9: ABH Secretor Phenotypes and HIV Infection Rates. There was a higher incidence rate of HIV among ABH secretors in comparison to non-secretors.

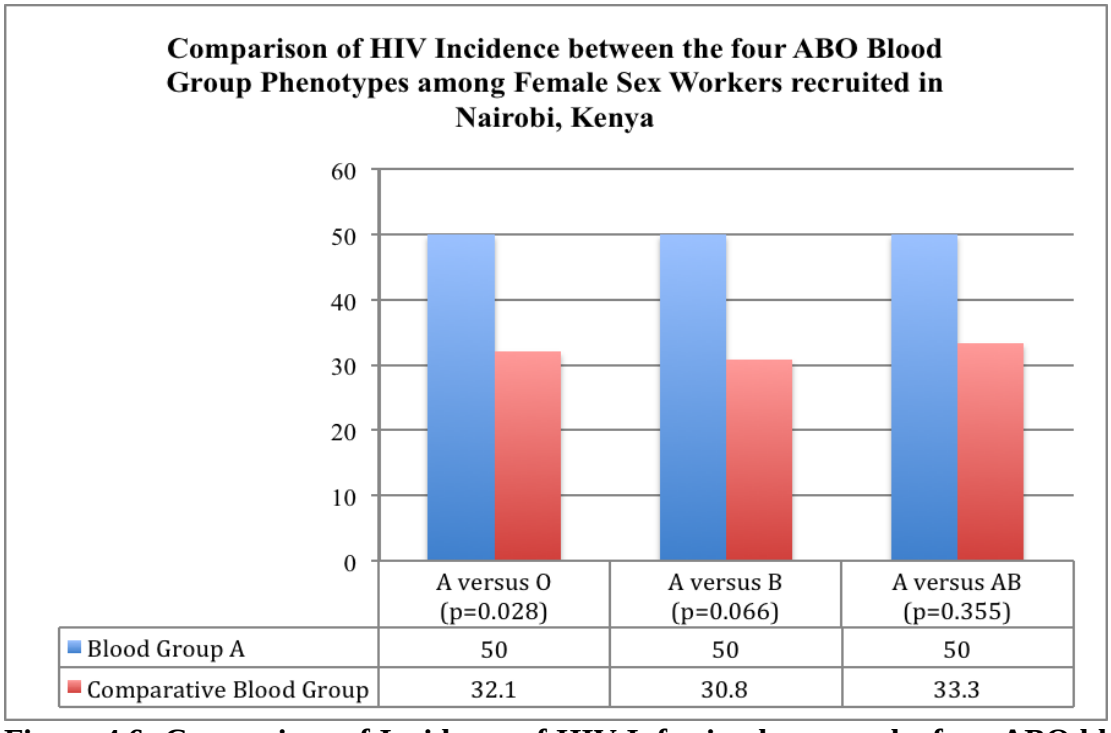


Figure 4.6: Comparison of Incidence of HIV Infection between the four ABO blood group phenotypes.

4.5.4 Vaginal/Cervical Infections:

All female sex workers underwent a detailed gynaecological examination and samples were screened for Bacterial Vaginosis (BV), Neisseria gonorrhea (GC) and Trichomonas vaginalis (TV) as indicated. For all STI cases that tested positive, the secretors were a majority however; there was no significant difference in the incidence of infection between secretors and non-secretors when comparing those who tested positive for BV, GC and TV (Table 4.10).

	n / No. Tested	**Secretors (%)**	**Non-secretors (%)**	**P Value**
Trichonomas (Positive)	11/166 (6.6%)	11 (6.6%)	0 (0%)	-
Trichonomas (Negative)	155/166 (93.4%)	123 (74.1%)	32 (19.3%)	
Neisseria gonorrheae (Positive)	69/167 (41.3%)	56 (33.5%)	13 (7.8%)	p=0.930
Neisseria gonorrheae (Negative)	98/167 (58.7%)	79 (47.3%)	19 (11.4%)	
Bacterial vaginalis (Positive)	78/174 (44.8%)	64 (36.7%)	14 (8%)	p=0.335
Bacterial vaginalis (Negative)	96/174 (55.2%)	73 (42%)	23 (13.3%)	

Table 4.10: *Bacterial vaginosis*, *Neisseria gonorrhea* and *Trichomonas vaginalis* infections among female sex workers recruited from the Pumwani Majengo cohort (n=280)

4.5.5 Confounding Factors

The association between HIV-1 and secretor status, as outlined above, was independent of several potentially confounding variables including, age, nationality, contraceptive use, and for *Bacterial vaginosis, Neisseria gonorrhea* and *Trichomonas vaginalis* infections.

4.6 Secretor Status, HIV and CD4+ T Cell Counts

To investigate correlations between secretor status and HIV infection based on CD4+ T cell counts as a surrogate marker of disease progression, secretor status, HIV sero-status and CD4+ T cell counts were determined for 90 female sex workers in Nairobi, Kenya.

4.6.1 HIV Screening:

Study participants were screened for HIV-1. 52 (58%) were HIV positive and 38 (42%) HIV negative (Figure 4.7).

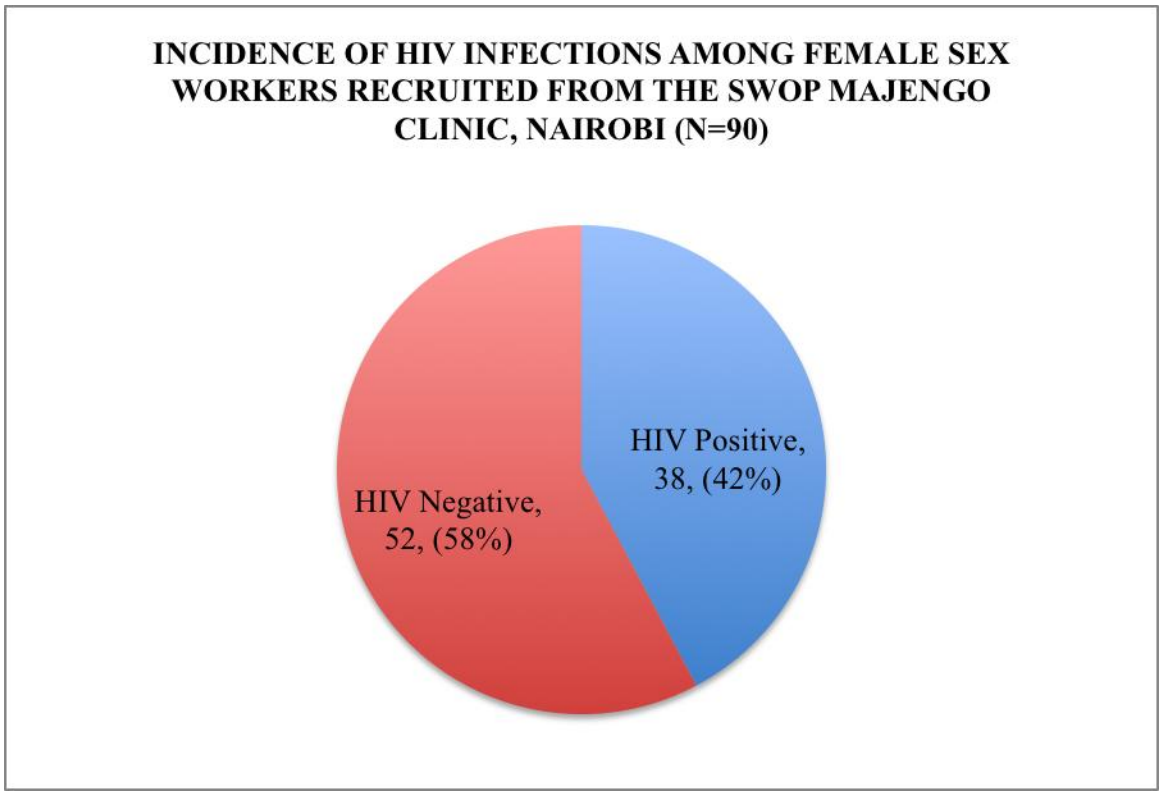


Figure 4.7: HIV Infection Rates in the among female sex workers recruited from the SWOP Majengo clinic (n=90)

4.6.2 Secretor Status:

Based on saliva screening 77 (85.6%) of the participants were secretors and 13 (14.4%) were non-secretors. This was consistent in swabs obtained from vaginal and cervical sampling sites (Figure 4.8).

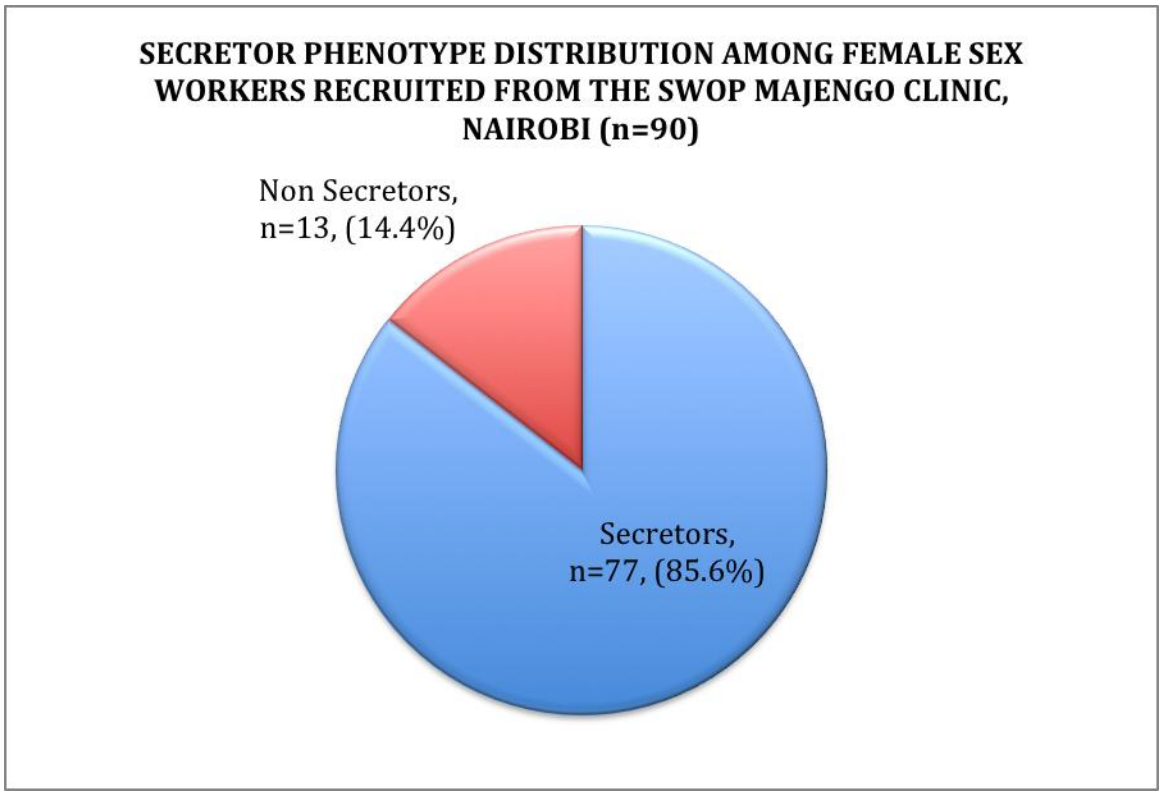


Figure 4.8: Secretor Phenotype Distribution among Female Sex Workers recruited from the SWOP Majengo clinic, Nairobi (n=90)

4.6.3 Secretor Phenotypes and HIV Infection Rates:

The proportion of secretors was significantly higher among women with HIV infection (36/38 = 94.7%) than among the HIV un-infected women (41/52 = 78.8%) p=0.034 (Table 4.11).

Phenotype	HIV Positive	HIV Negative	P Value
Secretors	36 (94.7%)	41 (78.8%)	p = 0.034
Non-secretors	2 (5.3%)	11 (21.2%)	
Totals	38	52	

Table 4.11: Distribution of secretor and non-secretor phenotypes among HIV infected and HIV uninfected female sex workers recruited from the SWOP Majengo clinic, Nairobi (n=90)

4.7 CD4 Cell Counts

4.7.1 Medical History:

Study participants were required to complete a health assessment questionnaire during their clinic visits, outlining their medical history over the past one year. Of these 23 of the 38 HIV infected women (60.5%) were on antiretroviral therapy, while 15 were ART naïve. This data was correlated with CD4 cell counts, and of those not on ART, 3 out of 15 had CD4+ T cell counts < 350 cells/mm^3 at the time of their clinic visit (Figure 4.9).

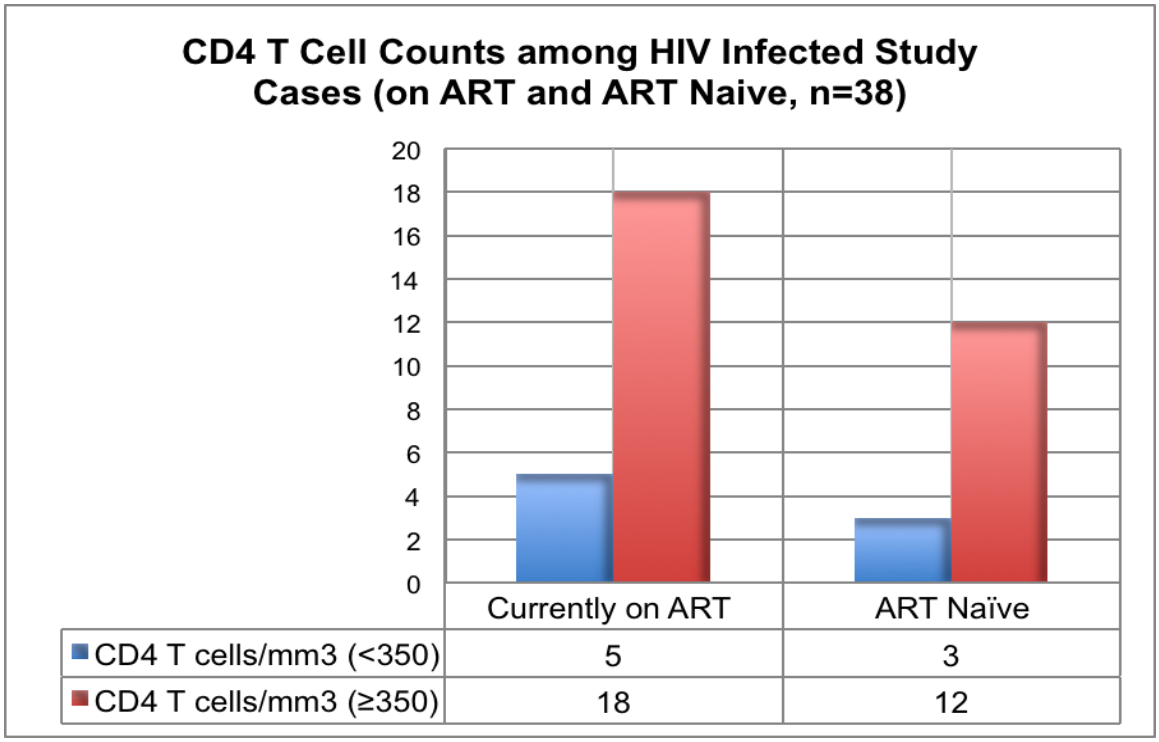


Figure 4.9: CD4 T cell counts among HIV Infected Study Cases. 60% of the women were on ART while 40% were ART naïve.

The HIV infected study cases were further stratified as secretors and non-secretors based on CD4 T cell counts (Table 4.12). The proportion of secretors with CD4 T cell counts ≥ 350 cell/mm^3 (96.8%) was higher in comparison to the non-secretors and this difference was statistically significant (p=0.005). The secretors, although more susceptible to HIV infections, maintained a significantly higher proportion of CD4 T cell counts.

Phenotype	CD4 T cell/mm^3 < 350	CD4 T cell/mm^3 ≥ 350	p Value
Secretors	6 (85.7%)	30 (96.8%)	p = 0.005
Non-secretors	1 (14.2%)	1 (3.2%)	
Totals	7	31	

Table 4.12: Stratification of secretor phenotypes against CD4 T cell counts.

4.8 Secretor Gene Polymorphisms

Secretor FUT2 genotype profiles among 66 female sex workers, all Kenyan nationals, aged 18-65 years were determined.

4.8.1 HIV Sero-status

Of the 66 women, 27 (41%) were HIV infected and 39 (59%) HIV un-infected as shown in Figure 4.10 below:

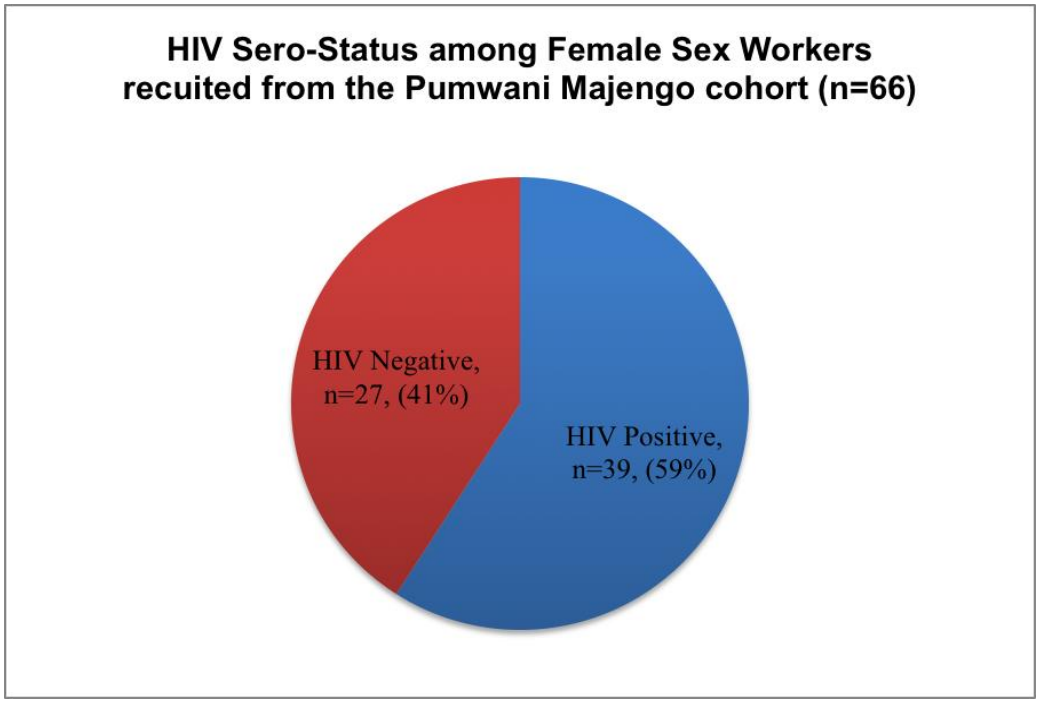


Figure 4.10: HIV sero-status among female sex workers recruited from the Pumwani Majengo cohort (n=66)

4.8.2 Secretor Phenotypes

The study cases comprised 8 (12%) secretors and 58 (88%) non-secretors. This was based on screening for the blood group H antigen in saliva, vaginal and cervical samples as shown in Figure 4.11 below:

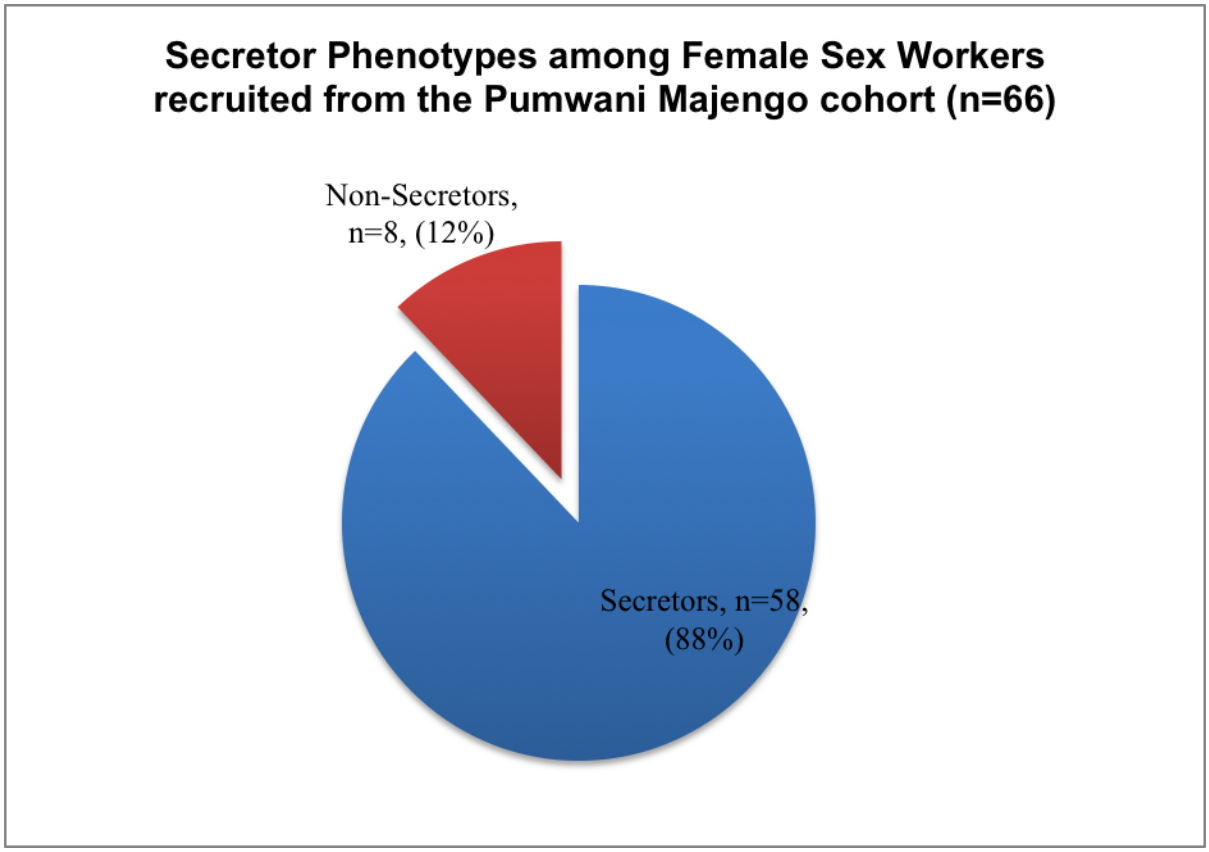


Figure 4.11: Secretor status among Female sex workers recruited from the Pumwani Majengo cohort (n=66)

4.8.3 Confirmation of PCR products:

The amplified PCR products were analysed by electrophoresis in an agarose/Tris EDTA gel containing ethidium bromide and visualised under UV light. After the PCR amplification, 20 µl of the amplified product was used for subsequent Pyrosequencing analysis as shown in Figure 4.12 below:

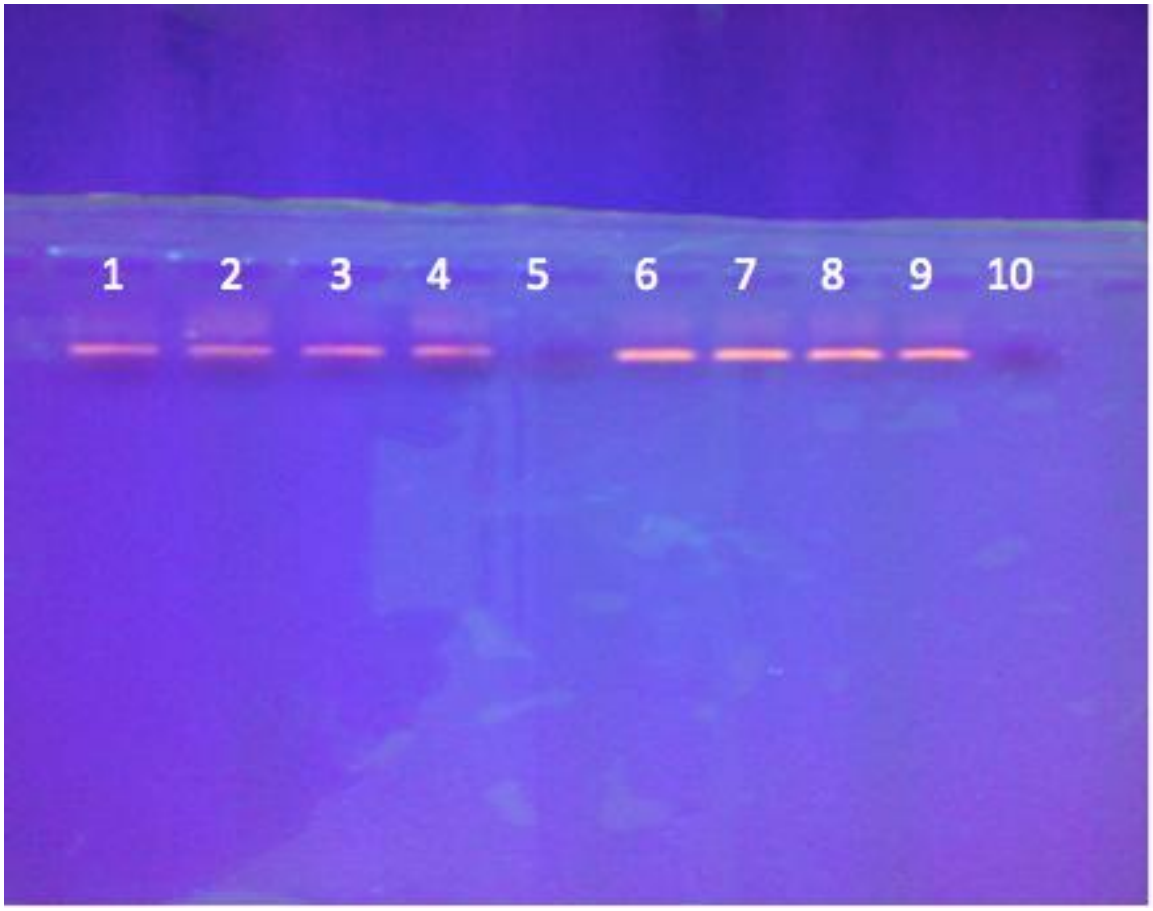


Figure 4.12: Gel image of amplified PCR products.
Ten wells are shown above, well 1-4 are samples with confirmed amplified products, well 5 is water (negative control); 6-9 samples with confirmed amplified products, well 10 is water (negative control)

4.8.4 Secretor Genotypes: Screening for the FUT2 se[428] null allele

To screen for the FUT2 se[428] null allele, sequences were mapped to the FUT2 reference. Both the mapped and reference sequences were analyzed for occurrence of restriction sites. The region of interest was identified to have SmaI restriction site (Table 4.13 and Figure 4.12) in both the reference and mapped reads.

Name	Location	HGVS* Name	Sequence
rs150802597	423	NM_000511.5:c.423	GCCACATCCC G GGGGAGTACG

Table 4.13: Restriction site in both reference and mapped reads
***HVGS - Human Genome Society Variation**

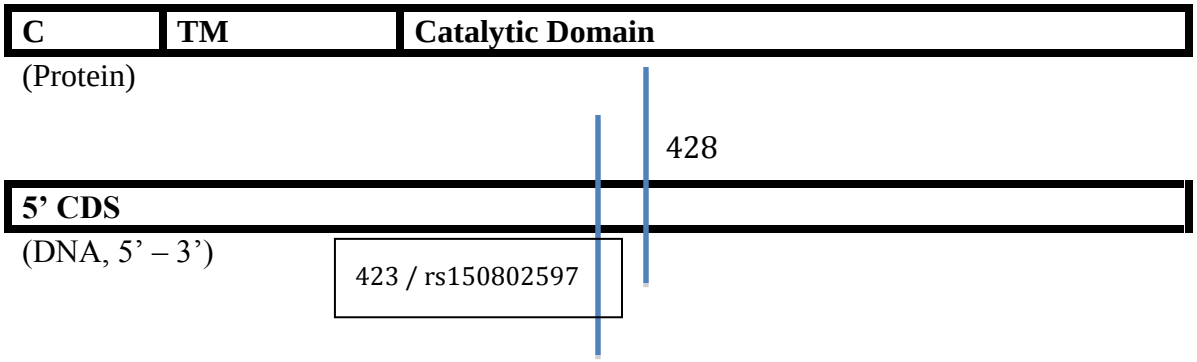


Figure 4.12: FUT2 Gene Sequencing site, screening at position 428.

The protein coding sequence (CDS), of FUT2 are shown in the Figure 12 above. The corresponding domains are organized in the NH2-terminal hydrophobic cytoplasmic domain 'C', a golgi-membrane spanning domain (Transmembrane, TM) and the COOH-terminal catalytic domain residing in the lumen of the golgi apparatus. All the reads to the FUT2 reference sequence were mapped with a focus at base 428. Observing the bases after position 428 there was a constant pattern of AGT, occurring after a GGGG sequence as shown in Table 4.14. This was for all 65-study cases.

Sample ID	420	421	422	423	424	425	426	427	428	429	430	431	432
Reference	C	C	C	G	G	G	G	G	G	**A**	**G**	**T**	A
#1	C	C	C	G	G	G	G	G	G	**A**	**G**	**T**	A
#27	C	C	C	G	G	G	G	G	G	**A**	**G**	**T**	A
#40	C	C	C	G	G	G	G	G	G	**A**	**G**	**T**	A
#45	C	C	C	G	G	G	G	G	G	**A**	**G**	**T**	A
#67	C	C	C	G	G	G	G	G	G	**A**	**G**	**T**	A

Table 4.14: Base arrangement on gene focus region comparing a subset of samples and the reference sequence. Samples #1, #27, #40 (Non secretors), #45, #67 (Secretors)

There were no sequence variations observed at position 428, based on the SmaI restriction site (rs150802597) at position 423 and followed by a reference sequence: *GGGGAGTACGTC *. These results demonstrate the secretor status of the population under study is not dependent on the allele at position 428 as has been previously reported.

5 DISCUSSION

This is the first study to date to profile, the variable expression of the blood group antigens (secretor status) in the Kenyan population. The study began with a baseline profile of the ABO, Rhesus (D) and Duffy blood group antigens. The prevalence patterns of these blood groups among the study population were similar to findings from previous studies in African populations. Notably, 1) predominance of blood group O witnessed in similar study populations 2) high prevalence of the Rhesus (D) positive phenotype (95.1%) similar to findings from an earlier study in Kenya 94%, Nigeria ranging from 93%-98% and 3) high prevalence of the erythroid Duffy null phenotype (Fy a-b-) (Loua *et al.*, 2007; Pennap *et al.*, 2011; Erbahor *et al.*, 2013; Mwangi, 1999; Egesie *et al.*, 2008; Ukaejiofor *et al.*, 1996; Pogo and Chaudhuri, 2000; Daniels *et al.*, 2005). These findings confirm, human populations globally share the same blood groups systems; the difference only comes in, in the frequencies of the specific phenotypes.

In addition, this is the first study on the frequency of secretor phenotypes: secretors *versus* non-secretors in this setting. The *fucosyltransferase 2 (FUT2)* gene, encodes for the enzyme alpha 1,2 fucosyltransferase which results in the expression of the H antigen and thereby A and B antigens, expressed in epithelial cells, body tissues and in mucosal secretions (Storry and Olsson, 2009). Phenotyping was therefore, determined by screening for the variable expression of the H antigen in saliva samples. The ABH secretors comprised 78.9% while the non-secretors a 21.1% minority among the target population. Globally, frequency distribution patterns of secretors vary markedly; 84.4% secretors and 15.6% non-secretors as reported in a study in Nigeria (Olorunshola and Audu, 2013), 64.4% secretors, 35.6% non-secretors, in Pakistan (Saboor *et al.*, 2014),

60% secretors, 40% non-secretors in Dhaka (Akhter *et al.*, 2011), 73% secretors, 26.9% non-secretors in Hungary (Karpati *et al.*, 2014), and 78.8% secretors, 21.2% non-secretors in Burkina Faso (Nordgren *et al.*, 2013).

This data fits into the global picture with non-secretors representing a minority, although the frequency rates are variable. The variability in prevalence rates may be linked to the underlying genetic profiles. Ferrer-Admetlla *et al.*, (2009) and Koda *et al.*, (2001), describe natural histories of the *Secretor FUT2* gene polymorphisms outlining a number of population-specific polymorphisms. Furthermore, heterozygous secretors (*Sese*) express decreased levels of the resultant protein (enzyme: alpha 1,2 fucoslytansferase) in comparison to their homozygous (*SeSe*) counterparts. Therefore for the heterozygotes, determination of secretor status based solely on phenotyping may result in an underestimation of the proportion of secretors in a population.

Questions have been raised on the physiological significance of blood group antigens, aside from their well-described role in blood transfusion and compatibility testing. Indeed, functional relevance of a number of these moieties has been experimented and documented (Lin *et al.*, 2009). There is the increasing interest in knowledge on the tissular distribution of the ABH antigens in humans, often in correlation to disease as evidenced by studies on the associations between secretor phenotypes and a number of mucosal infections including Noroviruses (Nordgren *et al*., 2013) and *Streptococcus pneumonia* (Higgins *et al*., 2009).

The findings of this study confirm the hypothesis; secretors are more susceptible to HIV infections as the results demonstrate a significantly higher proportion of secretors among HIV infected female sex workers in comparison to the HIV-uninfected ($p=0.029$). This

suggests the presence of the ABH blood group antigens in mucosal secretions may increase an individual's risk to HIV infection while the non-secretor phenotype, lack of expression of ABO antigens in mucosal secretions, may confer a certain degree of protection against HIV infection. Furthermore, when stratified based on the individual ABO phenotypes, blood group A secretors had a higher incidence of infection in comparison to their blood group AB, B and O counterparts. These findings compare with those of Ali *et al.*, (2000) who reported similar findings among Senegalese commercial sex workers, in which non-secretors were found to have a reduced HIV-1 seroprevalence. In a second study, Kindberg *et al.*, (2006) reported a slower HIV-1 disease progression in non-secretors.

Decades of research have established, HIV can cross the female genital tract epithelium through several routes; via transcytosis across an intact epithelium to the luminal surface; by penetration between the epithelial wall gaps, or if there are ulcerations or abrasion wounds, the virion gains direct access into the sub-mucosa (Bomsel, 1997; Bobardt *et al.*, 2007). This grants HIV access to the underlying target cells, after-which the cascade resulting in establishment of a systemic infection begins. In this case, the basis underlying the expression of blood group antigens in the genital mucosa and HIV infection may possibly be mediated by the selective binding to the A, B and H blood group antigens via a family of receptors called lectins, which bind specifically to carbohydrate moieties **(**Anstee, 2010**)**. It is postulated that these pathogens may initiate infection via non-covalent binding to these mucosal cell surface carbohydrate-binding proteins (lectins) (Arendrup *et al.*, 1991; Storry and Olsson, 2009;). In this way, expression of the blood group antigens on mucosal surfaces in ABH secretors may place

the host at an increased risk of infection (by aiding in pathogen attachment). On the contrary, modification of cell surface carbohydrates at mucosal surfaces in ABH non-secretors may in other instances offer a protective role.

A second mechanism may be based on the premise that various pathogens including bacteria and viruses express blood group antigen-identical or cross-reactive molecules on their surfaces (Higgins *et al.,* 2009). These pathogens have been postulated as probable targets for the corresponding blood group antibodies. Early studies on viral glycosylation (Arendrup *et al.*, 1991) supported this line of thought. HIV isolates cultured *in vitro* with peripheral blood mononuclear cells from donors of different ABO groups demonstrated neutralization by the corresponding ABO antibodies of specific cell isolates. Preece *et al.*, (2002) demonstrated that measles virus, when co-cultured in a system expressing the ABH glycosyltransferases *in vitro*, expressed the corresponding A or B epitopes, according to the enzymes expressed. In addition, Neil *et al.*, (2005) highlight that HIV-1 can incorporate ABO blood group antigens, both in an artificial trans-infection system, and when primary strains are propagated in human peripheral blood mononuclear cells.

In the context of HIV infections, CD4+ T cells are the hallmark of disease progression. CD4+ T lymphocyte numbers in the human body are kept constant by homeostatic mechanisms. However, during HIV infection, these mechanisms fail, and AIDS sets in characterized by progressive immune deficiency. Loss of CD4+ T cells and systemic immune hyper-activation are the hallmarks of acute HIV infection while chronic HIV infection is characterized by the massive production of pro-inflammatory cytokines, which further results in the gradual loss of peripheral CD4+ T cells over time (Haase, 2005). In relation to secretor status; the ABH secretors in this population were found to

have an increased susceptibility to HIV infection, however, following infection, the same ABH secretors maintained high levels of CD4+ T cell counts ≥ 350 cells/mm^3, the cut off for ART initiation at the time of the study.

This is interesting, because HIV research has been primarily focused on cells of the immune system, excluding research on viral interactions with red blood cells. However, emerging evidence suggests that red blood cells may play a significant role in the HIV infectious process by either binding free virions or binding to viral immune complexes thereby facilitating trans-infection to HIV target cells. These findings demonstrate correlations between the variable expression of the ABH blood group antigens (secretor status), T cells and HIV infection rates. This alone underscores the importance of widening the diagnostic scope in the context of HIV infections to include T cells and red blood cells, including secretor phenotypes.

Finally, single nucleotide polymorphisms (SNPs) are becoming increasingly important in the field of genomics. This study was designed to profile for the non-functional *se*^{428} null-allele, a SNP that has been associated with the non-secretor status in African populations (Kelly *et al*., 1995; Liu *et al*., 1998). The study was among a cohort of female sex workers with known secretor and non-secretor phenotypes. An optimized next generation sequencing assay protocol was designed to screen for the SNP among the population. These findings indicate, at position 428, both secretors and non-secretors had no polymorphism as has been previously reported. The G-A mutation at this point results in a stop codon and thereby, translation of a non-functional form of the protein. However, in this population, among both the secretors and non-secretors there was a consistent G-G, with no polymorphism. These findings therefore suggest, the underlying genotypic

variation resulting in the non-secretor phenotype in a subset of the Kenyan population as screened among this study cases does not fall under a the ‘ se^{428} umbrella’ of a restricted geographical distribution as has been previously reported (Koda *et al*., 2001; Ferrer-Admettla *et al,* 2009; Kindberg *et al*., 2006; Ali *et al*., 2000).

Furthermore, the variation in the phenotypes and associated genotypes suggest that although the prevalence of secretors and non-secretors in this population is similar to profiles globally, there is an existent inter-population and probably intra-population and inter-individual variability in the context of genetics. There are underlying novel polymorphisms that confer both the secretor and non-secretor phenotypes within this setting.

6 CONCLUSION

It was presumed for decades that blood group antigens were exclusively limited to transfusion and compatibility medicine. It is now apparent; these moieties are of clinical significance and may contribute to the provision of first line of defense against pathogens. This study demonstrates there is a correlation between secretor status, and HIV. Furthermore these results indicate that *FUT2* diversity patterns cannot be explained by neutrality and human demography; there are underlying novel polymorphisms that confer both the secretor and non-secretor phenotypes within this setting. The postulated underlying mechanism of protection is, these carbohydrate moieties possibly enhance viral binding and viral penetration, leading to establishment of infection, which may be particularly important at the initial stages of viral uptake into cells of the female genital tract.

7 SUMMARY OF RESEARCH FINDINGS AND RECOMMENDATIONS FOR THE FUTURE

This section summarizes and briefly discusses the study's research findings with useful insight on possible future directions of the research work.

7.1 **Objective 1:** To determine ABO, Rhesus and Duffy phenotype frequency profiles in the Kenyan population.

7.1.1 Summary of Research Findings

This chapter presents a profile of the ABO, Rhesus and Duffy phenotypes among voluntary blood donors. Blood group O was the most frequent, followed by A, B and AB the least common. The profiles are similar to previous reports globally among blood donors of African descent. The study findings also confirm the high prevalence of Rhesus positive phenotype and Duffy negative phenotypes both, which are characteristic of African populations. Overall the characteristics of the blood donors described, with >70% aged under 30 years and predominantly male are as described in a review by (Tagny *et al.*, 2009) describing the characteristics of blood donors in sub Saharan Africa.

7.1.2 Recommendations

It is important to know the frequencies of various blood group antigen phenotypes in a population. This information is important to confirm the prevalence of both major and minor blood group phenotypes in a population. There are a number of blood group phenotypes, which have been associated to varying disease/infection susceptibilities as seen in this study, blood group A individuals were at an increased risk for HIV infections.

7.2 Objective 2: To profile secretor and non-secretor phenotype frequencies in the Kenyan population.

7.2.1 Summary of Research Findings

This is the first report to date on the frequency of the secretor phenotypes: secretors *versus* non-secretors in the Kenyan population. Phenotyping was determined by screening for the variable expression of the H antigen in saliva samples; hence the frequency of the H antigen expression in the saliva (secretors): 78.9% and lack of H antigen expression in the saliva (non-secretors): 21.1% among this target population comprising voluntary blood donors and female sex workers.

7.2.2 Recommendations

It is important to profile secretor and non-secretor phenotypes in a population as this information may contribute to the comprehension of varying disease susceptibility patterns in a population. Pathogens are highly selective agents; therefore for microorganisms that preferentially bind to carbohydrate moieties, host cell surface molecules such as the ABO antigens can reveal patterns of selection in a given population.

7.3 **Objective 3:** To investigate correlations between secretor status and HIV infection

7.3.1 Summary of Research Findings

This study enlisted 280 female sex workers aged 18-65 years from the Pumwani Majengo cohort, Kenya. 32.9% of the study cases were HIV-1 infected and 67.1% HIV-1 uninfected. Based on blood group H antigen screening in vaginal and cervical samples, 75.7% study cases were secretors and 24.3% non-secretors. The proportion of secretors was significantly higher among women with HIV infection (p=0.029). Based on ABO phenotype stratification, the incidence of HIV infection was significantly higher among blood group A secretors.

7.3.2 Recommendations

Routine secretor status screening should be encouraged, especially in areas with increased HIV prevalence rates and among the most at risk populations who are known to display concentrated HIV epidemics. This information is necessary because the secretor phenotype is associated with increased susceptibility to HIV infections. There is however need for further exploration into factors underlying the protective effect conferred by the non-secretor phenotype.

7.4 Objective 4: To investigate correlations between secretor status and HIV infection based on CD4+ T cell counts as a surrogate marker of disease progression.

7.4.1 Summary of Research Findings

This objective describes a correlation between secretor phenotypes and HIV infection rates among 90 female sex workers aged 18-65 years from the Pumwani Majengo cohort. The study findings indicate 42.2% of the study cases were HIV-1 infected and 57.8% HIV-1 uninfected. Based on blood group H antigen screening in vaginal and cervical samples, 85.6% study cases were secretors and 14.4% non-secretors. The proportion of secretors was significantly higher among women with HIV infection ($p=0.034$). However, following infection, the proportion of patients with CD4+ T cell counts ≥ 350 was significantly higher among secretors when compared to non-secretors (96.8% versus 85.7%) ($p=0.005$).

7.4.2 Recommendations

Secretor phenotype profiles should be determined among HIV sero-positive individuals. This information will contribute towards the prediction of disease progression patterns; as seen in this study the ABH secretors were able to maintain elevated levels of CD4+ T cell counts even after infection; based on CD4+ T cell counts as markers of disease progression.

7.5 Objective 5: Profile for the presence of the null allele se^{428} across the *FUT2* gene among secretors and non-secretors in the Kenyan population.

7.5.1 Summary of Research Findings

This study sequenced the second exon of the *FUT2* gene on a set of samples from a HIV high-risk population of female sex workers. The most common SNP associated with the non-secretor phenotype in African populations is the non-functional se^{428}. These findings indicate, at position 428, both secretors and non-secretors had no polymorphism as has been reported in other populations. The mutation at this point results in a change from G-A resulting in a stop codon and thereby, translation of a non-functional form of the protein. However, in this population, among both the secretors and non-secretors there was a consistent G-G, with no polymorphism. These findings therefore suggest, the underlying genotypic variation resulting in the non-secretor phenotype in a subset of the Kenyan population as screened among this study cases does not fall under a the 'se^{428} umbrella' of a restricted geographical distribution as has been previously reported.

7.5.2 Recommendations

Laboratory screening for secretor status should incorporate a multipronged approach: phenotyping and genotyping. If the screening is strictly genotypic, whole genome sequencing should be performed initially on a subset of the population, and the specific novel polymorphisms identified, and utilized as a baseline for the adoption of genetic screening. Further investigations should be performed to confirm phenotypes and associated genotypes, as each population is unique.

8 REFERENCES

1. Akhter S, Kibiria GM, Akhter NR, Habibullah MM, Islam SMK, Zakariah M. ABO and Lewis blood grouping with ABH secretor and non-secretor status: a cross sectional study in Dhaka. Faridpur Med. Coll. J. 2011; 6(1): 38-40
2. Ali S, Niang AM, N'doye I, Critchlow CW, Hawes SE, Hill AV, Kiviat NB. Secretor Polymorphism and Human Immunodeficiency Virus Infection in Senegalese Women. J Infect Disease. 2000; 181: 737-739
3. Anstee D. The relationship between blood groups and disease. Blood. 2010; 115: 4635-4643
4. Arendrup M, Hansen JE, Clausen H, Nielsen C, Mathiesen LR, Nielsen JO. Antibody to histo-blood group A antigen neutralizes HIV produced by lymphocytes from blood group A donors but not from blood group B or O donors. AIDS. 1991; 5:441-444
5. Argyris EG, Pomerantz RJ HIV-1 Vif versus APOBEC3G: newly appreciated warriors in the ancient battle between virus and host. Trends in Microbiology 2004; 12 (4): 145-148
6. Arthos J, Cicala C, Martinelli E, Maclead K, Van Ryk D, Wei D, Xiao Z *et al*., HIV-1 envelope protein binds to and signals through integrin alpha4beta7, the gut mucosal homing receptor for peripheral T cells. Nat Immunol. 2008; 9(3):301-9.
7. Avent ND, Reid ME. The Rh blood group system: a review. Blood. 2000; 95(2): 375-387
8. Ayatollahi H, Rafatpanah H, Khayyami ME, Sayyadpour D, Ravarian M, Sadeghian MN, Izadi N, Khoob MK. Association between ABO and Rhesus

blood group systems among confirmed human T lymphotropic virus type 1-infected patients in Northeast Iran. AIDS Res Hum Retroviruses. 2008; 24: 1155-8.

9. Ayouba A, Mauclere P, Martin PM, Cuinn P, Mfoupouendoun J, Njinku B, Souquieres S and Simon F. HIV-1 group O infection in Cameroon, 1986 to 1998. Emerg Infect Dis. 2001; 7(3): 466-467.

10. Bandyopadhyay S, Zhan R, Chaudhuri A, Watabe M, Pai SK, Hirota S, Hosobe S, Tsukada T, Miura K, Takano Y, Saito K, Pauza ME, Hayashi S, Wang Y, Mohinta S, Mashimo T, Liizumi M, Furuta E, Watabe K. Interaction of KAI1 on tumor cells with DARC on vascular endothelium leads to metastasis suppression. Nat Med. 2006; 12: 933-938

11. Baral S, Beyrer C, Muessig K, Poteat T, Wirtz AL, Decker MR, Sherman SG and Kerrigan D. Burden of HIV among female sex workers in low-income and middle-income countries: A systematic review and meta-analysis. Lancet Infect Dis. 2012; 380: 367-77

12. Birney EJ, Stamatoyannopoulos A, Dutta A, et al. (313 co-authors). 2007. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature. 447:799–816.

13. Boasso A, Shearer GM: Chronic innate immune activation as a cause of HIV-1 immunopathogenesis. Clin Immunol 2008, 126:235-242.

14. Bobardt MD, Chatterji U, Selvarajah S, Van der Schueren B, David G, Kahn B, and Gallay PA. Cell- free human immunodeficiency virus type 1 transcytosis through primary genital epithelial cells. J Virol. 2007. 81(1): p. 395- 405.

15. Bomsel, M., Transcytosis of infectious human immunodeficiency virus across a tight human epithelial cell line barrier. Nature medicine, 1997. 3(1): p. 42-.- 7.

16. Brenchley JM, Schacker TW, Ruff LE, Price DA, Taylor JH, Beilman GJ, Nguyen PL, Khoruts A, Larson M, Haase AT, Douek DC: CD4+ T cell depletion during all stages of HIV disease occurs predominantly in the gastrointestinal tract. J Exp Med 2004, 200:749-759.

17. Buchbinder SP, et al. Efficacy assessment of a cell-mediated immunity HIV-1 vaccine (the Step Study): a double-blind, randomised, placebo-controlled, test-of-concept trial. Lancet. 2008;372:1881–1893

18. Casimiro DR, et al. Attenuation of simian immunodeficiency virus SIVmac239 infection by prophylactic immunization with dna and recombinant adenoviral vaccine vectors expressing Gag. J. Virol.2005;79:15547–15555Charneau *et al.*, 1994;

19. Checkley MA, Luttge BG, Freed EO. HIV-1 envelope glycoprotein biosynthesis, trafficking, and incorporation. J Mol Biol, 2011; 410**,** 582-608.

20. Chilcott J, Lloyd Jones M, Wight J, Forman K, Wray J, Beverly C. A review of the clinical effectiveness of routine anti-D prophylaxis for pregnancy women who are Rhesus (RhD) negative. Lonon: National Institute of Clinical Excellence, 2002.

21. Chown B. Amaemia from bleeding of the fetus into the mother's circulation. Lancet. 1954; 263 (6824):1213

22. Clapham PR, McKnight A. Cell surface receptors, virus entry and tropism of primate lentiviruses. 2002; 83 (Pt 8):1809-29

23. Collins LM, Warnock ND, Tosh DG, McInnes C, Everest D, Montgomery WI, Scantlebury M, Marks N *et al.*, Squirrelpox virus: Assessing prevalence, transmission and environmental degradation. Plos One. 2014; doi: 10.1371/journal.pone.0089521

24. Crowley, LV. An Introduction to Human Disease: Pathology and pathophysiology correlations [With Workbook]. Eighth Edition. Jones & Barlett Pub (ma). 2009.

25. Crowther CA, Philippa M, McBain RD. Anti-D administration in pregnancy for preventing Rhesus alloimmunisation. Cochrane Database of Systematic Reviews. 2013. Issue 2. Art. No. CD000020. DOI: 10.1002/14651858.CD000020.pub2

26. Daniels G and Bromillow I (eds). The ABO blood groups in essential guide to blood groups. Third Edition. John Wiley & Sons, Oxford. 2013; doi: 10.1002/9781118688915.ch3

27. Daniels G, Van Der Schoot M, Olsson M. Report of the First International Workshop on molecular blood group genotyping. J Transfus Med. 2005; 88 (2): 136–142.

28. Darbonne WC, Rice GC, Mohier MA, Apple T, Herbert CA, Valente AJ, Baker JB. Red blood cells are a sink for interleukin 8, a leukocyte chemotaxin. J Clin Invest. 1991; 88(4):1362-9

29. Darwazeh AM, Lamey PJ, Samaranayake LP, MacFarlane TW, Fisher BM, Macrury SM, MacCuish AC. The relationship between colonization, secretor status and in-vitro adhesion of Candida albicans to buccal epithelial cells from diabetics. J Med Microbiol. 1990; 33(1): 43-9.

30. Dawson TC, Lentsch AB, Wang Z, Cowhig JE, Rot A, Maeda N, Peiper Sc.

Exaggerated response to endotoxin in mice lacking the Duffy antigen/receptor for chemokines (DARC). Blood. 2000; 1; 96(5):1681-4

31. De Leys R, Vanderborght B, Vanden Haesevelde M, Heyndrickx L, van Geel A, Wauters C, Bernaerts R, Saman E, Nijs P, Willems B, *et al*. Isolation and partial characterization of an unusual human immunodeficiency retrovirus from two persons of west-central African origin. J Virol. 1990; 64: 1207–1216

32. Derdeyn CA, Decker JM, Bibollet-Ruche F, Mokili JL, Muldoon M, Denham SA, *et al.* Envelope-Constrained Neutralization-Sensitive HIV-1 After Heterosexual Transmission. Science. 2004; 303 (5666):2019-2022

33. Dieffenbach CW, Fauci AS. Thirty years of HIV and AIDS: Future challenges and opportunities. Ann Intern Med. 2011; 154 (11):766-771.

34. Dimmock N, Easton A, Leppard K. Introduction to modern virology, 6th Edition. Wiley-Blackwell. 2007.

35. Edgren G, Hjalgrim H, Rostgaard K, Norda R, Wikman A, Melbye M and Nyren O. Risk of gastric cancer and peptic ulcers in relation to ABO blood type: A cohort study. Am J Epidemiol. 2010; 172: 1280-1285

36. Egesie UG, Egesie OJ, Usar I, Johnbull TO. Distribution of ABO, Rhesus blood groups and haemoglobin electrophoresis among the undergraduate students of Niger Delta University Nigeria. Niger J Physiol Sci. 2008; 23: 5–8.

37. Erhabor O, Adias TC, Jeremiah ZA, Hart ML. Abnormal hemoglobin variants, ABO, and Rhesus blood group distribution among students in the Niger Delta of Nigeria. Pathol. Lab. Med. 2010; 2:41-46

38. Erhabor O, Isaac IZ, Saidu A, Ahmed HM, Abdulrahaman Y, Festus A,

Ikhuenbor DB, Iwueke IP and Adias TC. The distribution of ABO and Rhesus blood groups among residents of Gusau, Zamfara State, North Western Nigeria. Research and Reviews: journal of Medical and Health Sciences. 2 (4). P. 58-63 (2013)

39. Estes JD, Wietgrefe S, Schacker T, Southern P, Beilman G, Reilly C, Milush JM, Lifson JD, Sodora DL, Carlis JV, Haase AT: Simian immunodeficiency virus-induced lymphatic tissue fibrosis is mediated by transforming growth factor beta 1-positive regulatory T cells and begins in early infection. J Infect Dis. 2007; 195:551-561.

40. Ferrer-Admetlla A, Sikora M, Laayouni H, Esteve A, Roubinet F, Blancher A *et al.* A Natural History of FUT2 Polymorphim in Humans. Molecular Biology and Evolution. 2009; 26 (9):1993–2003

41. Fiebig EW, Wright DJ, Rawal BD, Garrett PE, Schumacher RT, Peddada L, Heldebrant C, Smith R, Conrad A, Kleinman SH, Busch M. Dynamics of HIV viremia and antibody seroconversion in plasma donors: implications for diagnosis and staging of primary HIV infection. AIDS. 2003; 17 (13): 1871-1879.

42. Flegel W. Molecular genetics and clinical applications for *RH*. Transfus Apher Sci. 2011; 44(1): 81-91.

43. Ford ES, Puroned CE, Sereti I. Immunopathogenesis of asymptomatic chronic HIV infection: The calm before the storm. Curr Opin HIV AIDS. 2009; 4(3):206-214.

44. Fra AM, LOcati M, Otero K, Sironi M, Signorelli P, Massardi ML, Gobbi M, Vecchi A, Sozzani S, Mantovani A. Cutting edge: Scavenging of inflammatory

CC chemokines by the promiscuous putatively silent chemokine receptor D6. J Immunol. 2003; doi.10.4049/jimmunol.170.5.2279

45. Freed, EO, Martin, MA. HIVs and their replication. In: Knipe, David M.; H., PM.; Griffin, Diane E.; Lamb, Robert A.; Martin, Malcolm A.; Roizman, Bernard; Straus, Stephen E., editors. Fields Virology. 5th Edition. Lippincott Williams & Wilkins; 2006.

46. Fry AE, Griffiths MJ, AuburnS, Diakite M, Forton JT *et al.*, (2008) Common variation in the ABO glycosyltransferease is associated with susceptibility to severe Plasmodium falciparum malaria. Hum Mol Genet. 17: 567-576 doi: 10.1093/hmg/ddm331

47. Gelmon L, Kenya P, Oguya F, Cheluget B and Haile G. Kenya HIV Prevention Response and Modes of Transmission Analysis. Nairobi: Kenya National AIDS Control Council. 2009.

48. Gibbs A, Hirbod T, Li Q, Bohman K, Ball TB, Plummer FA, Kaul R, Kimani J, Broliden K and Tjernlund. Presence of CD8+ T cells in the ectocervical mucosa correlates with genital viral shedding in HIV-infected women despite a low prevalence of HIV RNA-expressing cells in the tissue. J Immunol. 2014. 192 (8). 3947-3957

49. Giri PA, Yadav S, Parhar GS, Phalke DB. Frequency of ABO and Rhesus Blood Groups: A study from a rural tertiary care teaching hospital in India. Int J Biol Med Res. 2011; 2(4):988-990

50. Goonetilleke N, Liu MK, Salazar-Gonzalez JF, Ferrari G, Giorgi E, Ganusov VV, Keele BF, Learn GH. The first T cell response to transmitted/founder virus

contributes to the control of acute viremia in HIV-1 infection. J Exp Med. 2009. 206:1253-72

51. Gotte M, Li X, Wainberg MA. HIV-1 reverse transcription: A brief overview focused on structure-function relationships among molecules involved in initiation of the reaction. Arch. Biochem. Biophys. 1999; 365(2): 199-210

52. Gouws E, White PJ, Stover J and Brown T. Short term estimates of adult HIV incidence by mode of transmission: Kenya and Thailand as examples. 2006. Sex Transm Infect. 82, (Suppl. 3): iii51-iii55.

53. Gray RG, Wawer MJ, Brookmeyer R, Sewankambo NK, Serwadaa D, Wabwire-Mangen F, Lutalo T, Li X, vanCott T, Quinn TC. Probability of HIV-1 transmission per coital act in monogamous, heterosexual, HIV-1 discordant couples in Rakai, Uganda. Lancet; 2001. 357 (9263):1149-1153

54. Grez M, Dietrich U, Balfe P, von Briesen H, Maniar JK, Mahambre G, Delwart, EL, Mullins JI, Rubsamen-Waigmann H. Genetic analysis of human immunodeficiency virus type 1 and 2 (HIV-1 and HIV-2) mixed infections in India reveals a recent spread of HIV-1 and HIV-2 from a single ancestor for each of these viruses. J Virol. 1994; 68 (4). 2161-2168.

55. Grossman Z, Meier-Schellershelm M, Paul WE, Picker LJ. Pathogenesis of HIv infection: what the virus spares is as important as what it destroys. Nat. Med. 2006; 12: 289-295.

56. Gu L, Krendelchtchikova V, Kredelechtchtchikov A, Oster RA, Fujihashi K, Matthews QL. A recombinant adenovirus-based vector elicits a specific humoral immune response against the V3 loop of HIV-1 gp120 in mice through the

"Antigen Capsid-Incorporation" strategy. J Virol. 2014; 11:112

57. Guillon P, Clement M, Sebille V, Rivain JG, Chou CF, Ruvoen-Clouet N and Le Pendu J. Inhibition of the interaction between the SARS-CoV spike protein and its cellular receptor by anti-histo-blood group antibodies. Glycobiology. 2008; 18 (12). 1085-1093.

58. Gurtler LG, Hauser PH, Eberle J, von BrunnA, Knapp S, Zekeng L, Tsague JM, Kaptue L. A new subtype of human immunodeficiency virus type 1 (MVP-5180) from Cameroon. J Virol. 1994; 68(3):1581-1585

59. Gurunathan S, El Habib R, Bagylos L, Meric C, Plotkin S, Dodet B, Corey L, Tartaglia J. Use of predictive markers of HIV disease progression in vaccine trials. Vaccine. 2009; 27(14) 1997-2015

60. Haase AT. Targeting early infection to prevent HIV-1 mucosal transmission. Nature. 2010; 464:217-223

61. Haase AT. Perils at mucosal front lines for HIV and SIV and ther hosts. Nat. Rev. Immunol. 2005; 5:783-789

62. Hamed CT, Bollahi MA, Abdelhamid I, Med Mahmoud MA, BA B, Ghabert S et al. Frequencies and ethnic distribution of ABO and Rh(D) blood groups in Mauritiana: results of first nationwide study. Int J Immunogenet. 2012; 39:151-4.

63. Harrington PR, Lindesmith L, Yount B, Moe CL, Baric RS. Binding of Norwalk virus-like particles to ABH histo-blood group antigens is blocked by antisera from infected human volunteers or experimentally vaccinated mice. J Virol. 2002; 76:12335-43

64. Hayes RJ and Bennett S. Simple sample size calculations for cluster-randomized

trials. Int J. Epidemiol. 1999; 28(2): 319-326

65. He B, Qiao X, Klasse PJ, Chiu A, Chadburn A, Knowles DM, Moore JP, Cerutti A: HIV-1 envelope triggers polyclonal Ig class switch recombination through a CD40-independent mechanism involving BAFF and C-type lectin receptors. J Immunol 2006, 176:3931-3941.

66. Henry S, Mollicone R, Fernandez P, Samuelsson B, Oriol R, Larson G. 1996. Molecular basis for erythrocyte Le(afl bfl) and salivary ABH partial-secretor phenotypes: expression of a FUT2 secretor allele with an A/T mutation at nucleotide 385 correlates with reduced alpha(1,2) fucosyltransferase activity. Glycoconj J. 13:985–993.

67. Higgins MA, Abbott DW, Boulanger MJ, Boraston AB. Blood group antigen recognition by a solute-binding protein from serotype 3 strain of Streptococcus pneumonia. J Mol Biol. 2009. 388 (2):299-309 doi: 10.1016/j.jmb.2009.03.012

68. Hirbod T, Kimani J, Tjernlund A, Cheruiyot J, Petrova A, Ball TB, Mugo N, Jaoko W, Plummer FA, Kaul R and Broliden K. Stable CD4 expression and local immune activation in the ectocervical mucosa of HIV-infected women. J Immunol. 2013. 191 (7):3948-3954

69. Hladik, F, McElrath, MJ, Setting the stage: host invasion by HIV. Nat Rev Immunol. 2008. 8(6):447-57

70. Horuk R, Chitnis CE, Darbonne WC, Colby TJ, Rybicki A, Hadley TJ, Miller LH. Science. 1993; 261(5125): 1182-1184

71. Howes RE, Patil AP, Piel FB, Nyangiri OA, Kabaria CW et al. The global distribution of the Duffy blood group. Nat Commun. 2011; 2:266

72. Hu Liya, Crawford SE, Czako R, Cortes-Penfield NW, Smith DF, Le Pendu J, Estes MK and Prasad BVV. Nature. 2012; 485:256-259.

73. Huang J, Kang BH, Pancera M, Lee HJ, Tong T, Feng Y, Georgiev IS, Chuang G, Druz A, *et al.*, Broad and potent HIV-1 neutralization by a human antibody that binds the gp41-gp120 interface. Nature. 2014; doi:10.1038/nature13601

74. Huang P, Farkas T, Zhong W, Tan M, Thornton S, Morrow AL, et al. Norovirus and histo-blood group antigens: demonstration of a wide spectrum of strain specificities and classification of two major binding groups among multiple binding patterns. J Virol. 2005; 79:6714-22.

75. Huston AM, Atmar RL, Graham DY, Estes MK. Norwalk virus infection and disease is associated with ABO histo-blood group type. J Infect Dis 2002; 185:1335-7

76. International Society of Blood Transfusion (ISBT, 2014) http://www.isbtweb.org/working-parties/red-cell-immunogenetics-and-blood-group-terminology/blood-group-terminology/ Accessed 2 November 2014: 3:53pm

77. Ishitoya S, Yamamoto S, Mitsumori K, Ogawa O, Terai A. Non-secretor status is associated with female acute uncomplicated pyelonephritis. BJU Int. 2002; 89 (9):851-854

78. Jamieson T, Cook DN, Nibbs RJ, Rot A, Nixon C, Mclean P, Alcami A, Lira SA, Wlekowski M, Graham GJ. The chemokine receptor D6 limits the inflammatory response *in vivo*. Nat. Immunol. 2005; 6:403-411

79. Jin J, Colin P, Staropoli I, Lima-Fernandes E, Ferret C, Demir A, Rogee S,

Hartley O *et al*., J. Biol. Chem. 2014; doi: 10.1074/jbc.M114.559831

80. Joint United Nations Programme on HIV/AIDS (UNAIDS). Global Report: UNAIDS report on the global AIDS epidemic 2013. (2013).

81. Kalayanarooj S, Gibbons RV, Vaughn D, et al. Blood group AB is associated with increased risk for severe dengue disease in secondary infections. J Infect Dis 2007; 195:1014-7.

82. Karlsson Hedestam GB, Fouchier RAM, Phogat S, Burton DR, Sodroski J, Wyatt RT. The challenges of eliciting neutralizing antibodies to HIV-1 and to influenza virus. Nat. Rev. Micro. 2008; 6:143-155.

83. Karpati K, Braunitzer G, Toldi J, Turzo K, Virag K, Reiche WT, Rakonczay Z, Nagy K. Caries and ABO secretor status in a Hungarian population of children and adolescents: an exploratory study. Caries Res. 2014; 48:179-185.

84. Keele BF, Giorgi EE, Salazar-Gonzalez JF, Decker JM, Pham KT, Salazar MG, Sun C, Grayson T *et al.* Identification and characterization of transmitted and early founder virus envelopes in primary HIV-1 infection. PNAS. 2008; 105(21):7552-7557

85. Kelly RJ, Rouquier S, Giorgi D, Lennon GG, Lowe JB. Sequence and expression of a candidate for the human Secretor blood group alpha(1,2)fucosyltransferase gene (FUT2). Homozygosity for an enzyme-inactivating nonsense mutation commonly correlates with the non-secretor phenotype. J Biol Chem. 1995; 270:4640–4649.

86. Kerrigan D, Moreno L, Rosario S, Gomez B, Jerez H, Barrington C, Weiss E, Sweat M. Environmental-structural interventions to reduce HIV/STI risk among

female sex workers in the Dominican Republic. American Journal of Public Health. 2010; 96 (1): 120-125

87. KinaneDF, Blackwell CC, Brettle RP, Weir DM, Winstanley FP, Elton RA. ABO blood group, secretor state, and susceptibility yo recurrent urinary tract infection in women. Br Med J. 1982; 285(6334):7-9

88. Kindberg E, Hejdeman B, Bratt G, Wahren B, Lindblom B, Hinkula J, Svensson L. A nonsense mutation (428G/A) in the fucosyltransferase FUT2 gene affects the progression of HIV-1 infection. AIDS. 2006; 20:685–689.

89. Koda Y, Tachida H, Pang H, Liu Y, Soejima M, Ghaderi AA et al. Contrasting patterns of polymorphisms at the ABO-secretor gene (FUT2) and plasma alpha (1, 3) fucosyltransferase (FUT6) in human populations. Genetics, 2001; 158 (2):747 – 756.

90. Lajoie J, Kimani M, Plummer FA, Nyamiobo F, Kaul R, Kimani J and Fowke KR. Association of sex work with reduced activation of the mucosal immune system. J Infect Dis. 2014. doi: 10.1093/infdis/jiu023

91. Landsteiner, K. Zur Kenntniss der antifermentativen, lytischen und agglutinierenden Wirkungen des Blutserums und der Lymphe. Zbl Bakt. 1900; 27: 357-366.

92. Larsson MM, Rydell GE, Grahn A, Rodriguez-Diaz J, Akerlind B, Hutson AM, Estes MK, Larson G, Svensson L. 2006. Antibody prevalence and titer to norovirus (genogroup II) correlate with secretor (FUT2) but not with ABO phenotype or Lewis (FUT3) genotype. J Infect Dis. 194:1422–1427.

93. Lasky LA, Nakamura G, Smith DH, Fennie C, Shimaski C, Patzer E, Berman P,

Gregory T, Capon DJ. Cell. 1987; 50(6):975-985

94. Le Pendu J, Ruvoen-Clouet N, Kindberg E, Svensson L. Mendelian resistance to human norovirus infections. Semin Immunol 2006;18: 375-86.

95. Levine P, Burnham L, Katzin WM, Vogel P. The role of isoimmunization in the pathogenesis of erythroblastosis. Am J Obset Gynecol. 1941; 42:925-37.

96. Li MJ, Kim J, Li S, Zaia J, Yee JK, Anderson J, Akkina R, Rossi JJ. Long-term Inhibition of HIV-1 infection in primary haematopoetic cells by lentiviral vector devlivery of a triple combination of anti-HIV shRNA, anti-CCR5 Robozyme and a nucleolar-localizing TAR decoy. Molecular Therapy. 2005; 12:900-909

97. Li Q, Estes JD, Schlievert PM, Duan L, Brosnahan AJ, Southern PJ, Reilly CS *et al.* Glycerol monolaurate prevents mucosal SIV transmission. Nature. 2009; 458:1034-1038.

98. Lialiaris T, Digkas E, Kareli D, Pouliliou S, Asimakopoulos D, Pagonopoulou O, Simopolou D. Distribution of ABO and Rh blood groups in Greece: An update. Int J Immunogenet. 2011; 38:1-5

99. Lin Y, Pavenski K, Saidenberg E and Branch DR. Blood group antigens and normal red cell physiology: A Canadian Blood Services Research and Development Symposium. 2009. Transfusion Medicine Reviews. 23 (4), p.292-309.

100. Linden S, Mahdavi J, Semino-Mora C, Olsen C, Carlstedt I, Boren T, et al. Role of ABO secretor status in mucosal innate immunity and H pylori infection. PLoS Pathog 2008;4:e2.

101. Lindesmith L, Moe C, Marionneau S *et al.*, Human susceptibility and resistance to

Norwalk virus infection. Nat Med 2003; 9: 548-53.

102. Little SJ, McLean AR, Spina CA, Richman DD, Havlir DV. Viral dynamics of acute HIV-1 infection. J Cell. Biol.1999; 190(6) 841-850.

103. Liu Y, Koda Y, Soejima M, et al. (11 co-authors). 1998. Extensive polymorphism of the FUT2 gene in an African (Xhosa) population of South Africa. Hum Genet. 103:204–210.

104. Loua A, Lamah MR, Haba NY, Camara M. Frequency of blood groups ABO and rhesus D in the Guinean population. Transfus Clin Biol. 2007; 14(5): 435-439.

105. Lu Z, Berson JF, Chen Y, Turner JD, Zhang T, Sharron M, Jenks MH, Wang Z, Kim J *et al.*, Evolution of HIV-1 coreceptot usage through interactions with distinct CCR5 and CXCR4 domains. PNAS.1997; 94 (12), 6426-6431

106. Marionneau S, Airaud F, Bovin NV, Le Pendu J, Ruvoen- Clouet N. Influence of the combined ABO, FUT2, and FUT3 polymorphism on susceptibility to Norwalk virus attachment. J Infect Dis. 2005 192:1071–1077.

107. Mattapallil JJ, Douek DC, Hill B, Nishimura Y, Martin M, Roederer M. Massive infection and loss of memory CD4+ T cell in multiple tissues during acute SIV infection. Nature. 2005; 434: 1093-1097

108. McBurney SP, Ross TM. Viral sequence diversity: challenges for AIDS vaccine designs. Expert Rev, 2008: 1405-1417

109. McDougal JS, Nicholson JK, Cross GD, Cort SP, Kennedy MS, Mawle AC. Bidning of the human retrovirus HTLV-III/LAV/ARV/HIV to the CD4 (T4) molecule: conformation dependence, epitope mapping, antibody inhibition, and potential for idotypic mimicry. J Immunol. 1986; 137 (9):2937-2944

110. McElrath MJ, De Rosa SC, Moodie Z, Durbey S, Kierstead L, Janes H, Defawe OD, Carter DK, Hural J *et al.*, HIV-1 vaccine induced immunity in the test-of-concept Step Study: a case-cohort analysis. Lancet, 2008; 372(9653):1894-1905

111. McMichael AJ and Rowland-Jones SL. Cellular immune responses to HIV. Nature, 2001: 410: 980-987

112. McMichael AJ, Borrow P, Tomaras GD, Goonetilleke N, Haynes BF. The immune response during acute HIV-1 infection:clues for vaccine development. Nat. Rev. Immunol. 2010; 10:11-23

113. Meo SA, Abdo AA, Sanie FM, Baksh ND, Al-Qhatani A, Shaikh ZA, Al-Drees AM. Transmission of hepatitis B virus through salivary blood group antigens in saliva. Journal of the College of Physicians and Surgeons Pakistan. 2010; 20 (7):444-448

114. Mestecky J, Wei Q, Alexander R, Raska M, Novak J, Moldoveanu Z. Humoral immune responses to HIV in the mucosal secretions and sera of HIV-infected women. Americal Journal of Reproductive Immunology. 2014; 71 (6):600-607

115. Moir S, Malaspina A, Li Y, Chun TW, Lowe T, Adelsberger J, Baseler M, Ehler LA, Liu S, Davey RT Jr, Mican JA, Fauci AS: B cells of HIV-1-infected patients bind virions through CD21-complement interactions and transmit infectious virus to activated T cells. J Exp Med 2000, 192:637-646.

116. Morgan WT, Watkins WM. The inhibition of the haemagglutins in plant seeds by human blood group substances and simple sugars. Br J Exp Pathol. 1953; 34:94-103

117. Morrow A, Meinzen-Derr, Huang P, *et al*. Fucosyltransferase 2 Non-Secretor and

Low Secretor status predicts severe outcomes in premature infants. Journal of Pediatrics, 2001; 158: 745 – 751.

118. Mwangi J. Blood group distribution in an urban population of patient targeted blood donors. East Afr Med J. 1999; 76(11): 615-8

119. Nance ST, Lomas-Francis C. Where are we in efforts to unravel the complexity of Rh to guide transfusion decisions? Transfusion. 2013; 53(11): 2840-2843

120. National AIDS and STI Control Programme, Ministry of Health, Kenya. September 2013. Kenya AIDS Indicator Survey 2012: Preliminary Report. Nairobi, Kenya.

121. National AIDS Control Council and National AIDS and STI Control Programme, Ministry of Health, Kenya. June 2014. Kenya HIV Prevention Revolution Road Map.

122. Neil S, McKnight A, Gustafsson K and Weiss R. (2005). HIV-1 incorporates ABO histo-blood group antigens that sensitize virions to complement-mediated inactivation. Blood. 2005; 105 (12): 4693 – 4699.

123. Neote K, Darbonne W, Ogez J, Horuk R, Schall TJ et al. Identification of a promiscuous inflammatory peptide receptor on the surface of red blood cells. J Biol Chem. 1993; 268: 12247-12249

124. Noizat-Pirenne F. Relevance of RH variants in transfusion of sickle cell patients. ISBT Science Series. 2012; 7(1):134-137

125. Nordgren J, Nitiema LW, Ouermi D, Simpore J, Svensson L. Host genetic factors affects susceptibility to Norovirus infections in Burkina Faso. 2013; PLoS ONE 8(7): e69557. doi: 10.1371/journal.pone.0069557

126. Olorunshola KV, Audu L. ABO (H) secretor status of sickle cell disease patients in Zaria, Kaduna State, Nigeria. Niger J Physiol Sci. 2013; 28:29-34

127. Palmisano L and Vella S. A brief history of antiretroviral therapy of HIV infection: success and challenges. Annali dell'Istituto Superiore di Sanità, 2011: 47 (1), p. 44-48

128. Pang H, Koda Y, Soejima M *et al*., Polymorphism of the human ABO-Secretor locus (FUT2) in four populations in Asia: indication of distinct Asian subpopulations. Annals of Human Genetics, 2001: 65 (5), p.429-437

129. Patnaik SK, Helmberg W, Blumenfeld OO. BGMUT: NCBI dbRBC database of allelic variations of genes encoding antigens of blood group systems. *Nucleic Acids Res.* 2012 Jan;40(1):D1023-9

130. Pennap GR, Envoh E and Igbawua. Frequency distribution of hemoglobin variants, ABO, and Rhesus blood groups among students of African descent. British Microbiology Research Journal. 1 (2): 33-40 2011.

131. Pitisuttithum P, et al. Randomized, double-blind, placebo-controlled efficacy trial of a bivalent recombinant glycoprotein 120 HIV-1 vaccine among injection drug users in Bangkok, Thailand. J. Infect. Dis. 2006;194:1661–1671.

132. Pogo A, Chaudhuri A. The Duffy protein: a malarial and chemokine receptor. Seminars in Hematology. 2000; 37 (2): 122–129.

133. Preece A, Strahan K, Devitt J, et al. (2002). Expression of ABO or related antigenic carbohydrates on viral envelopes leads to neutralization in the presence of serum containing specific natural antibodies and complement. Blood, 99; 2477 – 2482.

134. Pruenster M, Rot A. Throwing light on DARC. Biochemical Society Transactions, 2006: **34** (6), 1005 – 1008.

135. Rahman M, Al-Emram A, Saha SK, Ali LM, Bachar SC, Ahmed Farzana. Distribution of ABO and Rhesus Blood Groups among the population of Mymensingh District and the students of University of Dhaka, Bangladesh. Ibrahim Card Med J. 2011; 1(1):33-5

136. Raza M, Blackwell C, Molyneaux P, *et al*. Association between secretor status and respiratory viral illness. British Medical Journal, 1991: 303 (6806), 815 – 818.

137. Rerks-Ngarm S, et al. Vaccination with ALVAC and AIDSVAX to prevent HIV-1 infection in Thailand. N. Engl. J. Med. 2009

138. Rockx BHG, Vennema H, Hoebe CJPA, Duizer E and Koopmans MPG. Association of Histo-Blood group antigens and susceptibility to Norovirus infections. J Infect Dis 2005; 191:749-54

139. Ronchetti F, Villa MP, Ronchetti R et al. ABO/Secretor genetic complex and susceptibility to asthma in childhood. Eur Respir J 2001; 17: 1236–1238.

140. Rowland-Jones S and Whittle H. Out of Africa: what can we learn from HIV-2 about protective immunity to HIV-1? Nature Immunology, 2007: 8, 329 - 331

141. Ruiz-Palacios GM, Cervantes LE, Ramos P, Chavez-Munguia B, Newburg DS. Campylobacter jejuni binds intestinal H(O) antigen (Fuc alpha 1,2Gal beta 1, 4GlcNAc), and fucosyloligosaccharides of human milk inhibit its binding and infection. *J Biol Chem* 2003: 278:14112-20.

142. Saboor M, Ullah A, Qamar K, Mir A, Moinuddin. Frequency of ABH secretors

and non secretors: A cross sectional study in Karachi. Park J Med Sci. 2014; 30 (1). P.189-193.

143. Salazar-Gonzalez J, Salazar M, Keele B *et al*., Genetic identity, biological phenotype, and evolutionary pathways of transmitted/founder viruses in acute and early HIV-1 infection. Journal of Experimental Medicine, 2009; 206 (6), 1273-1289

144. Schiff F, Sasaki H. Der Aus-scheidungstypus, ein auf serologischem wege nachweisbares Mendelndes Merkmal. *Klin Wschr*. 1932; 11: 1426.

145. Segurel L, Thompson EE, Flutre T, Lovstad J, Venkat A, Margulis SW, Moyse J, Ross S, Gamble K, Sella G, Ober C, Przeworski M. The ABO blood group is a trans-species polymorphism in primates. PNAS. 2012; 109(45): 18493-18498

146. Seltsam A, Hallensleben M, Kollmann A, Blasczyk R. The nature of diversity and diversification at the ABO locus. Blood. 2003; 102(8): http://dx.doi.org/10.1182/blood-2003-03-0955

147. Shen H, Schuster R, Stringer K *et al*., The Duffy antigen/receptor for chemokines (DARC) regulates prostate tumor growth. The Journal of the Federation of American Societies for Experimental Biology, 2006; 20(1), p.59-64

148. Slomiany BL, Slomiany A. Blood-group-(A+H) complex fucolipids of hog gastric mucosa. Eur J Biochem. 1978; 90: 39–49.

149. Smith JD, Cooper JD, Howson JMM, Clarke P, Downes K, Mistry T, Stevens H, Walter NM, Todd JA. FUT2 Nonsecretor status links type 1 diabetes susceptibility and resistance to infection. Diabetes. 60(11):3081-84.

150. Soejima M, Pang H, Koda Y. Genetic variation of FUT2 in a Ghanaian

population: identification of four novel mutations and inference of balancing selection. Ann Hematol. 2007; 86:199-204

151. Stapleton A, Nudelman E, Clausen H, Hakomori S, Stamm WE. Binding of uropathogenic *Escherichia coli* R45 to glycolipids extracted from vaginal epithelial cells is dependent on histo-blood group secretor status. *J Clin Invest.* 1992; 90: 965-72

152. Storry JR, Olsson ML. The ABO blood group system revisited: a review and update. *Immunohaematology*, 2009; **25** (2): 48-59.

153. Stowell SR, Arthur CM, Dias-Baruffi M, Rodrigues LC, Gourdine J-P, Heimburg-Molinaro J, Ju T, Molinaro RJ, Rivera-Marrero C, Xia B, Smith DF, Cummings RD. Innate immune lectins kill bacteria expressing blood group antigen. (2010). *Nature*., 16 (3): 295-301 doi: 10.1038/nm.2103

154. Swanstrom R. Wills, JW. Synthesis, assembly, and processing of viral proteins.. In: Coffin, JM. Hughes, SH. Varmus, HE, editors. Retroviruses. Cold Spring Harbor Laboratory Press; Cold Spring Harbor, NY: 1997

155. Tan M, Jiang X. Norovirus-host interactions: Multi-selections by human histo-blood group antigens. (2011). *Trends in Microbiology*., 19: 382-388.

156. Teixeira VP, Martins E, Almeida HO, Soares S, de Souza HM, de Morais CA. The ABO system and anatomoclinical forms of chronic Chagas disease. Rev Soc Bras Med Trop. 1987; 20:163-7.

157. Temitayo OI, Timothy SO. Frequency distribution of hemoglobin variants and rhesus blood groups among pregnant women. Am. Med. J. 2013; 4:78-81

158. Thorven M, Grahn A, Hedlund KO, Johansson H, Wahlfrid C, Larson G,

Svensson L. 2005. A homozygous nonsense mutation (428G/A) in the human secretor (FUT2) gene provides resistance to symptomatic norovirus (GGII) infections. J Virol. 79:15351–15355.

159. Ueno F, Shiota H, Miyaura M, Yoshida A, Sakurai A, Tatsuki J, Koyama AH, Akari H, Adachi A, Fujita M Vpx and Vpr proteins of HIV-2 up-regulate the viral infectivity by a distinct mechanism in lymphocytic cells. Microbes and Infection 5 (5): 387-395 APR 2003

160. Ukaejiofor EO, Okonkwo WC, Tagbar EN, Emeribe AO. Blood Transfusion in the Tropics. Nigeria: (Ukaejiofor EO) Salem Press; 1996. ABO and Rhesus in a Nigerian population; pp. 1–22.

161. Von Decastella A, Sturli A. Ureber die iso agglutinine in serum gesunder and Kranaker Menschen. Mfiner Med WSchr. 1902; 49:1090-5

162. Watkins WM, Morgan WT. Inhibition by simple sugars of enzymes which decompose the blood-group substance. Nature. 1955; 175:676-7

163. Wei X, Decker J, Wang S *et al*., (2003). Antibody neutralization and escape by HIV-1. Nature, **422**, p.307-312

164. Weijing H, Stuart N, Hemant K *et al.*, (2008). Duffy Antigen Receptor for Chemokines Mediates trans-Infection of HIV-1 from Red Blood Cells to Target Cells and Affects HIV-AIDS susceptibility. Cell Host & Microbe, **4** (1), p. 52 – 62.

165. Wolofsky KT, Ayi K, Branch DR, Hult AK, Olsson ML, Liles CW, Cserti-Gazdewich CM and Kain KC. ABO blood groups influence macrophage-medicated phagocytosis of Plasmodium falciparum-infected erythrocytes. Plos

Pathog 8(10): e1002942. doi: 10.1371/journal.ppat.1002942

166. Wolpin BM, Chan AT, Hartge P, Chanock SJ, Kraft P, Hunter DJ, Giovannucci EL, Fuchs CS. ABO blood group and the risk of pancreatic cancer. J Natl Cancer Inst. 2009; 101: 424-31

167. Wood LF, Chahroudi A, Chen HL, Jaspan HB, Sodora DL. The oral mucosa immune environment and oral transmission of HIV/SIV. Immunol Rev. 2013; 254:34-53

168. Yamamoto F, Clausen H, White T, Marken J, Hakomori S. Molecular genetic basis of the histo-blood group ABO system. Nature. 1990a; 345:229-33

169. Yamamoto F, Marken J, Tsuji T, Whote T, Clausen H, Hakomori S. Cloning and characterization of DNA complementary to human UDP-GalNAc:Fuc-alpha1-2Gal alpha1-3GalNAc transferase (histo-blood group A transferase) mRNA. J Biol Chem 1990b; 265:1146-51

170. Yu LC, Yang YH, Broadberry RE, Chen YH, Chan YS, Lin M. 1995. Correlation of a missense mutation in the human Secretor alpha 1,2-fucosyltransferase gene with the Lewis(aflbfl) phenotype: a potential molecular basis for the weak Secretor allele (Sew). Biochem J. 312 (Pt 2):329–332.

171. Zhang X-F, Tan M, Chhabra M, Dai Y-C, Meller J and Jiang X. Inhibition of histo-blood group antigen binding as a novel strategy to block norovirus infections. (2013), *Plos One*, 8(7): e69379. Doi:10.1371/journal.pone.0069379.

172. APPENDICES

8.1 Ethics Approval

8.2 Participant Information Forms, Consent Forms and Questionnaires